Document AO1
Appl. No. 09/987,763

PCT WORLD I ZATION

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 :
C12N 15/00, C12P 13/08
C12N 9/04, 9/12, 9/88

A2

(11) International Publication Number: WO 88/ 09819

(43) International Publication Date: 15 December 1988 (15.12.88)

(21) International Application Number: PCT/US88/02029

(22) International Filing Date: 10 June 1988 (10.06.88)

(31) Priority Application Number: 062,552

(32) Priority Date: 12 June 1987 (12.06.87)

(33) Priority Country: US

(71) Applicant: MASSACHUSETTS INSTITUTE OF TECHNOLOGY [US/US]; 77 Massachusetts Avenue, Cambridge, MA 02139 (US).

(72) Inventors: SINSKEY, Anthony, J. ; 285 Commonwealth Avenue, Boston, MA 02115 (US). FOLLETTIE, Maximillian, T. ; 474 Memorial Drive, Cambridge, MA 02139 (US). LIEBL, Wolfgang ; Hart Str. 17, D-8057 Eching (DE). PEOPLES, Oliver, P. ; 104 Fairmont Street, Arlington, MA 02174 (US).

(74) Agent: PABST, Patrea, L.; Kilpatrick & Cody, 100 Peachtree Street, Suite 3100, Atlanta, GA 30043 (US).

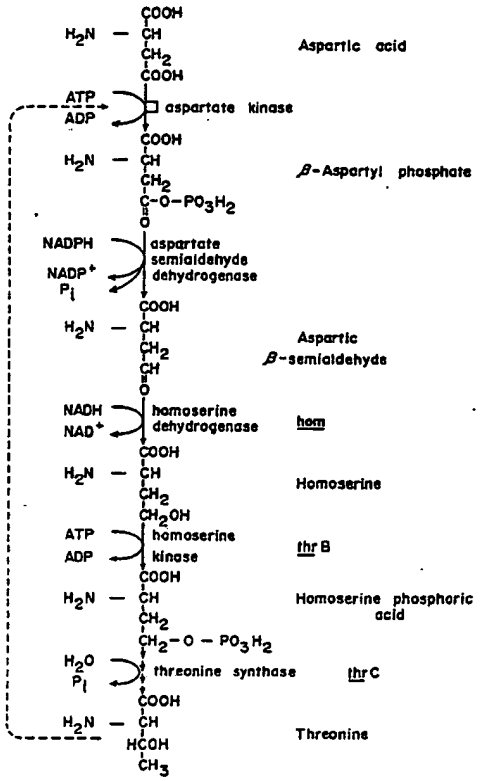
(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: C. GLUTAMICUM THREONINE BIOSYNTHETIC PATHWAY

(57) Abstract

The present invention is a method for the isolation and characterization of *C. glutamicum* genes involved in amino acid biosynthesis, specifically, encoding *hom*, *thr*B, and *thr*C, and sequences regulating their expression. Techniques for modifying or replacing these sequences and means for facilitating further isolations and characterizations, including promotor probe vectors which are useful in screening for high efficiency and regulatable promoters and repressors, are also disclosed. A *C. glutamicum* genomic library was constructed by cleaving chromosomal DNA with restriction enzymes, inserting the DNA fragments into an appropriate vector, and transforming the resulting recombinant molecules (rDNA) into *C. glutamicum*. Amino acid biosynthetic genes *hom*, *thr*B, and *thr*C, encoding homoserine dehydrogenase, homoserine kinase, and threonine synthetase, respectively, were isolated by complementation of *C. glutamicum* auxotrophs. The *hom-thr*B genes were subcloned on a 3.6 kb *Sal*I generated chromosomal fragment while *thr*C activity was isolated from a second recombinant plasmid within the genomic library and subcloned on a 2.7 kb *Sph*I generated fragment. The *hom-thr*B and *thr*C loci, and regulatory sequences, were identified by enzyme assays, complementation of defined *E. coli* auxotrophs, S1 nuclease and deletion mapping.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**ML**	Mali
AU	Australia	**GA**	Gabon	**MR**	Mauritania
BB	Barbados	**GB**	United Kingdom	**MW**	Malawi
BE	Belgium	**HU**	Hungary	**NL**	Netherlands
BG	Bulgaria	**IT**	Italy	**NO**	Norway
BJ	Benin	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		

C. GLUTAMICUM THREONINE BIOSYNTHETIC PATHWAY

Background of the Invention

The present invention is generally in the field of genetic engineering, and specifically, in the area of manipulation of amino acid biosynthesis in Gram positive bacteria.

Corynebacterium glutamicum is a Gram positive, nonpathogenic microorganism that has long occupied a central role in the industrial production of amino acids by conventional fermentation processes. Past strain development has primarily depended on classical mutagenesis to remove competing pathways to thereby increase substrate availability, and to remove or reduce regulatory control of a particular biosynthetic pathway. Regulatory mutants were isolated by selecting strains resistant to toxic amino acid analogues. The use of chemical mutagenesis has been very successful and a number of economically viable L-amino acid fermentation strains, such as strain producing L-glutamate and L-lysine, have been established.

The recent development of cloning vectors, including those described in U.S. Patent No. 4,649,119 to Sinskey et al., and methods for DNA transformation of C. glutamicum, as decribed by Katsumata et al., J. Bacteriol. 159,306-311 (1984), and Yoshihama et al., J. Bacteriol. 162, 591-597 (1985), and the closely related Corynebacterium (Brevibacterium) lactofermentum described by

Santamaria et al. in J. Gen. Microbiol. 130, 2237-2246 (1984), initiated a new era in the genetic manipulation of these organisms.

However, the commercial utilization of C. glutamicum recombinant DNA technologies for future strain development is dependent on the development of additional genetic tools and a better understanding of the fundamental molecular biology of this species. The use of recombinant DNA techniques to develop industrial strains would offer several advantages over classical mutagenic strategies. For example, specific alterations such as the replacement of a low efficiency promoter would be possible, the stepwise isolation of enhancing mutations could be avoided, regulatory systems could be engineered to allow the temporal control of gene expression during a fermentation process, and novel genes and/or pathways could be introduced into an organism.

It is therefore an object of the present invention to isolate and characterize genes encoding components of amino acid biosynthetic pathways in Corynebacterium.

It is another object of the present invention to clone the isolated amino acid biosynthetic genes, specifically those involved in the threonine biosynthetic pathway.

It is still another object of the present invention to elucidate the structure of these genes

and the regulatory mechanisms that modulate their expression.

It is a further object of the present invention to characterize and modify the expression of the cloned, amino acid biosynthetic genes, as well as the primary structure and regulatory features of their protein products.

Summary of the Invention

The present invention is a method for the isolation and characterization of C. glutamicum genes involved in amino acid biosynthesis, specifically, hom, thrB, and thrC, and sequences regulating their expression. Techniques for modifying their expression and regulation are also described. Methods and sequences facilitating further isolations and characterizations are also disclosed, including promoter probe vectors which are useful in screening for high efficiency and regulated promoters.

A C. glutamicum genomic library was constructed by cleaving chromosomal DNA with the restriction enzyme MboI, inserting the resultant DNA fragments into a C. glutamicum/Bacillus subtilis shuttle vector, pHY416, and transforming the resulting recombinant molecules into C. glutamicum. Amino acid biosynthetic genes hom, thrB, and thrC, encoding homoserine dehydrogenase, homoserine kinase, and threonine synthase, respectively, were isolated by complementation of C. glutamicum auxotrophs. The

hom-thrB genes were subcloned on a 3.6 kb Sal1 generated chromosomal fragment while thrC activity was isolated from a second recombinant plasmid within the genomic library and subcloned on a 2.7 kb Sph1 generated fragment. The hom-thrB and thrC loci were identified by a combination of enzyme assays and complementation of defined E. coli auxotrophs, and amino acid sequence homology.

Enzymatic assay of homoserine dehydrogenase activity, encoded by hom, in strains harboring the cloned gene demonstrated a 20-fold increase in specific activity compared to wild type controls. Both the chromosomal and plasmid encoded activities are strongly inhibited by L-threonine and repressed by L-methionine. The L-methionine repression of the plasmid encoded activity demonstrates that the structural gene and sequences responsible for its expression are included within the cloned fragment. Southern hybridization analysis demonstrated that the hom/thrB and thrC loci are separated by a minimum of 8.8 kb in the C. glutamicum chromosome. This is a different genomic organization from that observed in E. coli where the three genes represent a single operon. Three lines of evidence demonstrate that the C. glutamicum hom-thrB genes represent an operon. First, they are located together (separated by 11 base pairs) and coordinately regulated by L-methionine. Secondly, Northern hybridization analysis has identified a single 2.4 kb, L-methionine repressed RNA transcript,

consistent with the size of the two coding regions. Finally, deletion of the promoter upstream of the hom gene significantly reduces the expression of both the hom and thrB genes.

The hom-thrB and thrC promoters were identified by complementation of auxotrophs, deletion analysis and S1 nuclease mapping. The hom-thrB operator, a hyphenated dyad symmetry element, was also identified by deletion analysis. Methods for modification, removal or replacement of these regulatory elements are described.

Brief Description of the Drawings

Figure 1 is a schematic of the threonine biosynthetic pathway.

Figure 2 is a graphic depiction of subcloning strategy and restriction maps of recombinant plasmids pFS78, pFS80, pFS3.6A, pFS3.6B, pSPC1 and pSPC4.

Figure 3 is the nucleotide sequence and predicted protein sequences of hom and thrB.

Figure 4 is the nucleotide sequence and predicted protein sequence of the thrC gene.

Figure 5a is the sequence of the C. glutamicum hom-thrB regulatory region indicating the mRNA initiation site, -35 and -10 regions of thrP1 and the hyphenated dyad symmetry element responsible for methionine mediated repression (thrO).

Figure 5b is the potential stem/loop structure formed by the hyphenated dyad symmetry element.

Figure 6 is a schematic of the construction of C. glutamicum hom-thrB promoter deletions and subsequent analysis.

Figure 7: Deletional analysis of the hom-thrB promoter region. The 3.6 kb C. glutamicum chromosomal DNA insert of pFS3.6 carrying the hom-thrB genes is indicated as a hatched box and the nucleotide sequence of the relevant promoter-containing DraI-HindIII fragment is shown. The extent of Bal31 generated deletions in various plasmid constructs based on vector pWST1 are presented as black bars. The start of transcription as determined by S1 nuclease mapping is indicated by an arrow.

Detailed Description of the Invention

Recombinant DNA technology has been used to isolate, characterize and manipulate genes involved in the amino acid biosynthetic pathway of Corynebacterium glutamicum. The technology and results obtained aid in the elucidation of the fundamental molecular biology of C. glutamicum and construction of amino acid producing strains, particularly threonine.

Threonine is produced in a series of reactions beginning with the reduction of the beta-carboxyl group of aspartic acid to from the aldehyde, aspartic beta-semialdehyde, which takes place via an acyl phosphate intermediate, beta-aspartyl phosphate, in an ATP requiring reaction. Aspartic

beta-semialdehyde is converted by homoserine dehydrogenase, encoded by hom, to homoserine. Homoserine is phosphorylated by homoserine kinase, encoded by thrB, to homoserine phosphate in an ATP requiring reaction. The product, homoserine phosphoric acid, is in turn converted to threonine by threonine synthase, encoded by thrC, a pyridoxal phosphate enzyme. Threonine, the end product of the sequence, is an inhibitory modulator of aspartate kinase. This reaction pathway is demonstrated in Figure 1.

Genes encoding the three enzymes, homoserine dehydrogenase (hom), homoserine kinase (thrB), and threonine synthetase (thrC) have been isolated, identified, cloned, and their expression modified as follows. The cloning and determination of the nucleotide sequence of these genes provides a means for manipulating the expression and catalytic properties of the encoded enzymes. Means for altering the expression and the end product include in vitro mutagenesis of the C. glutamicum hom gene and selection of derivatives resistant to L-threonine mediated feedback inhibition, sequence determination of feedback resistant derivatives and the use of rDNA techniques to combine separate genetic alterations, determination and modification of the promoter structure and protein start sites for hom, thrB, and thrC, increased expression of hom-thrB via increased promoter efficiency and removal of L-methionine transcriptional repression,

and molecular joining of the C. glutamicum hom-thrB and thrC genes to form a hom-thrBC operon.

Isolation of thr- and thr-/met- auxotrophs of C. glutamicum.

A genetic background in which to isolate the threonine biosynthetic genes was constructed by mutating C. glutamicum and isolating auxotrophs defective in threonine biosynthesis. The C. glutamicum is maintained on LB media (10 g NaCl, 10 g Bactotryptone, 5 g Yeast extract, 1 l. H_2O) or minimal medium for C. glutamicum (MCG) (10 g glucose, 7 g $(NH_4)_2SO_4$, 3g K_2HPO_4, 1g KH_2PO_4, 0.4 g $MgSO_4.7\ H_2O$, 2 mg $FeSO_4 \cdot {}^7H_2O$, 2 mg $MnSO_4.H_2O$, 1 mg Biotin, 10 mg Thiamine, 2 ml trace elements, 1 l.H_2O). 1.4% agar was added for plates. The trace elements contained 44 mg $Na_2B_4O_7,.7H_2O$, 20 mg $(NH_4)_6Mo_7O_{27}\cdot 4H_2O$, 5 mg $ZnSO_4$, 135 mg $Cu50_4.5H_2O$, 3.6 mg $MnCl_2.H_2O$, 435 mg $FeCl_3$ in 500 ml H_2O. Where appropriate, 50 g/ml L-threonine, 50 g/ml L-methionine, 50 g/ml ampicillin, 15 g/ml kanamycin or 10 g/ml rifampicin were added.

C. glutamicum AS019, a rifampicin resistant variant of ATCC 13059, was grown at 30°C in LB to exponential phase (2 x 10^8 cfu/ml), harvested by centrifugation and resuspended in an equal volume of minimal media for C. glutamicum (MCG). Cells were mutagenized by the addition of nitrosoguanidine (NTG) (40 micrograms/ml) to 1 ml of cells and incubation without shaking at 30°C for 30 minutes. Mutagenized cells were harvested by centrifugation,

resuspended in 1 ml LB media and diluted 1:100 into 10 ml aliquots of fresh LB. Following growth at 30°C with shaking to stationary phase, the cells were diluted and plated on LB agar. Auxotrophs were screened by replica plating onto MCG plates and identification by growth patterns on amino acid pools. Only one strain displaying a particular auxotrophy such as threonine requirement was saved from each of the 10 ml aliquots.

Twenty four thr- and six thr$^-$/met- C. glutamicum auxotrophs were isolated. The thr/met- auxotrophs grow on MCG plates supplemented with homoserine. The thr- auxotrophs may have mutations in either of the threonine specific enzymes, homoserine kinase or threonine synthase.

Transformation and complementation of the C. glutamicum auxotrophs.

Two threonine requiring auxotrophs of C. glutamicum, AS155 and AS178, were transformed using the following method. An overnight culture of AS019 was inocculated at a ratio of 1:100 into LB broth containing 0.2% glucose and 2.0% glycine. The cells were incubated at 30°C for 15 hours with aeration. 10 ml of cells were harvested by centrifugation and washed in SMMC buffer (0.5/M Sorbitol, 20 mM $MgSO_4$, 20 mM $CaCl_2$, 50 mM Na Maleate, pH 7.0). Cells were resuspended in 2 ml SMMC buffer containing 2.5 mg lysozyme/ml. The cell suspension was incubated at 37°C with shaking for 90 minutes. Cells were again harvested by centrifugation at 6000 rpm for ten

minutes and resuspended in three ml SMMC buffer. 0.3 ml aliquots of "protoplasted" cells were placed in polypropylene tubes. Plamid DNA in 0.5 M sorbitol was added. 0.7 ml of 40% PEG, molecular weight 3350, 50 mM Tris, 20 mM $CaCl_2$ pH 7.4 was added and gently mixed. 2.0 ml of SB broth (0.5 M sorbitol, 1 x LB, 20 mM $CaCl_2$, 20 mM $MgSO_4$) was added to the transformation mixture, which was then incubated at 30°C without shaking for three hours. The C. glutamicum protoplasts obtained by growth in glycine and lysozyme treatment can also be suspended in SMMC and frozen at -80°C for use in subsequent transformations.

The transformants were plated out on selected plates. The two threonine requiring auxotrophs AS155 and AS178 were transformed with a C. glutamicum genomic library containing approximately 2.5 genomic equivalents constructed in the C. glutamicum/B. subtilis chimeric plasmid pHY416, described by Yoshihama et al., J. Bacteriol. 162, 591-597 (1985) and Follettie and Sinskey in J. Bacteriol. 166 695-702 (1986). Kanamycin resistant transformants were selected and screened for complementation of the threonine auxotrophy by replica plating onto MCG/Km plates. Three AS155 transformants and a single AS178 transformant were capable of growth without threonine supplementation. Plasmids were isolated and characterized by restriction analysis.

All four transformants harbored the same recombinant plasmid, designated pFS78, described in Figure 2, which contain a 6.8 kb chromosomal DNA insert. The recombinant plasmid was transformed into 10 of the independently isolated C. glutamicum threonine auxotrophs and three auxotrophs requiring threonine and methionine or homoserine supplementation. The Km^r transformants were screened for complementation on MCG/Km plates. The results demonstrated that pFS78 complements all three homoserine auxotrophs and four of the ten thr- auxotrophs, indicating that the plasmid carries the homoserine dehydrogenase gene, hom, as well as one of the threonine specific genes, thrB or thrC.

Two of the thr- auxotrophs not complemented by pFS78, AS148 and AS213, were transformed with the genomic library and Km^r colonies screened for growth on MCG/Km plates. Both thr+ AS148 and thr+ AS213 transformants were obtained, and their plasmids isolated and characterized by restriction analysis. All thr+ transformants harbor the same 12.5 kb recombinant plasmid designated pFS80, also shown in Figure 2, containing a 3.1 kb chromosomal DNA insert. The chromosomal sequence cloned in pFS80 complements four other thr- auxotrophs not complemented by pFS78. However, pFS80 was unable to complement the thr- or thr-/met- strains complemented by pFS78.

Subcloning and identification by enzyme assay, complementation of auxotrophs, and amino acid

sequence homology of the C. glutamicum hom, thrB, and thrC locus.

Deletion analysis of pFS78 indicated that both the thr- and thr-/met- complementing activities are located on a 3.6 kb Sal1 generated chromosomal fragment. This fragment was purified by agarose gel electrophoresis and electroelution and ligated into the unique Sal1 restriction site of the C. glutamicum/E. coli chimeric vector pWS124, described by Batt, Shamnabruch and Sinskey, Biotech. Letts. 7:717 (1985). The recombinant vector, pFS3.6, complements both AS178 (thr-) and AS253 (thr-/met-). The plasmid also complements E. coli thrB auxotroph, E. coli 5076.

The 2.7 Kb Sph1 generated chromosomal fragment of pFS80 was purified by agarose gel electrophoresis and ligated into the unique Sph1 restriction site of pUC18. The resulting recombinant plasmids, designated pSPC1 and pSPC4, also diagrammed in Figure 2, were able to complement E. coli 5077 (thrC) but not E. coli 5076 (thrB).

Southern hybridization analysis was used to determine the relationship of the hom-thrB and thrC loci. The results demonstrate that the hom-thrB and thrC locus in this species are physically separated by a minimum of 8.8 kb.

The homoserine dehydrogenase activity in crude extracts of wild type AS019 was compared to that of the homoserine auxotroph AS253 with and without the complementing plasmid pFS3.6 in order to determine

the regulation and extent of overproduction of the cloned C. glutamicum hom gene product. Methods used for preparing a C. glutamicum crude extract preparation and assays for homoserine dehydrogenase, homoserine kinase and aspartokinase are as follows:

Homoserine dehydrogenase is measured by the decrease in absorbance at 340 nm due to the oxidation of NADPH (extinction coefficient = 6220). The reaction mixture contains: 3 mM DL aspartate-beta-semialdehyde (ASA), 0.4 mM NADPH, 0.1 M PO_4, pH 7.0, 0.5 M KCl, and enzyme preparation, in a total volume of 0.7 ml. A blank reaction mixture without DL-ASA serves as a control. DL-ASA is synthesized by the ozonolysis of DL-allyl glycine according to the procedure of Black and Wright, J. Biol. Chem. 213, 39 (1955).

Homoserine kinase activity was determined by a coupled enzyme assay which measured the reaction product ADP. The reaction mixture contained 3.3 mM ATP, 0.45 MM NADH, 4.5 mM phophenol pyruvate, 1.0 mM L-homoserine, 10 mM $MgCl_2$ 12.5 units pyruvate kinase (Sigma, St. Louis, MO), 25 units lactate dehydrogenase (Sigma), 0.25 M KCl, 100 mM HEPES buffer (pH 7.8) and enzyme preparation in a total volume of 1.0 ml. The reaction was monitored by the decrease in absorbance at 340 nm due to the oxidation of NADH. The absorbance decrease in the absence of added substrate, L-homoserine, was determined and subtracted from values obtained with the complete assay mixture.

Aspartate kinase activity, inhibited by threonine, is determined by measuring the aspartohydroxamate produced according to the procedure of Black and Wright, J. Biol. Chem. 213, 27 (1955). Protein in the crude extract is precipitated by adding 5 volumes of saturated ammonium sulfate and resuspended in 0.3 volume of buffer containing 0.1 M Tris, pH 7.4, 0.2 M KCl. The assay mixture contains: 0.1 M Tris, pH 7.4, 10 mM ATP, 10 mM $MgSO_4$ 0.6 M hydroxylamine (pH 7.4), 0.6 M $(NH_4)_2$, 50 mM L-aspartate and enzyme preparation in a total volume of 1 ml. After 1 hr incubation at 37°C, the reaction was stopped by the addition of 1.5 ml of solution containing 10% $FeCl_3.6\ H_2O$, 3.3% trichloroacetic acid and 0.7 N HCl. After centrifugation, aspartohydroxamate concentration is measured by absorption at 540 nm (extinction coefficient = 600). A blank reaction mixture without L-aspartic acid serves as a control.

Protein concentration of the crude extracts is determined using the Bio-Rad protein assay with bovine serum albumin standards (BioRad Laboratories, Richmond, CA).

The C. glutamicum thr^-/met^- strain AS253 harboring the parental vector pWS124 had less than 2.5% of the homoserine dehydrogenase activity present in the wild type AS019. Introduction of the cloned C. glutamicum hom gene present on pFS3.6A into C. glutamicum AS253 leads to a twenty-fold increase in the specific activity of homoserine

dehydrogenase over that observed in wild type C. glutamicum AS019. The orientation of the cloned hom gene with respect to the vector affected its expression. Crude extracts of AS253 harboring pFS3.6B demonstrated an 11-fold increase in homoserine dehydrogenase activity relative to wild type.

The level of aspartokinase in C. glutamicum AS019 harboring either the parental vector pWS124 or the recombinant vector pFS3.6 was unchanged over that observed in the controls. Further, the aspartokinase specific activity was not repressed by growth in MCG supplemented with 2.7 mM L-methionine. The differential transcriptional control of homoserine dehydrogenase, in combination with the lack of increased aspartokinase activity in cells harboring pFS3.6 (hom-thrB), demonstrates that the two activities are not catalyzed by a bifunctional protein as in E. coli. The expression of the encoded homoserine dehydrogenase is repressed 3.2 fold by the addition of 2.7 mM L-methionine. Expression of the C. glutamicum thrA gene is also repressed by L-methionine, demonstrating that the expression of the pFS3.6 encoded hom gene is mediated by its native promoter/operator. Expression of the cloned C. glutamicum thrB gene was similarly repressed 2.6-fold by 2.7 mM L-methionine.

The activity of the homoserine dehydrogenase, both chromosomal and plasmid encoded, is inhibited by the addition of L-threonine to the assay mixture.

Addition of 1 mM D-threonine or L-methionine does not affect the homoserine dehydrogenase activity.

The complete nucleotide sequence of 3704 bp is shown in Figure 3 for hom and thrB. Two long open reading frames (ORF's) extend from position 907 to 2329 and from 2312 to 3269. The protein sequences of homoserine dehydrogenase and homoserine kinase are predicted on the basis of the sequence extending from the first potential translation initiation codon, either ATG or GTG (position 994) to the TAA stop at position 2330 for ORF1 and from the ATG at 2342 to the TAG stop 3269. The predicted proteins have molecular weights of 46,436 and 32,618 daltons for ORF1 and ORF2, respectively. A translation terminator is present at position 3279 to 3311, seven nucleotides downstream of the TAG stop codon. This is shown in further detail in Figure 5. The sequence forms a strong step-loop structure having a stem length of 15 bp and a seven base loop similar to the rho-independent terminators from E. coli. The 5' sequence to ORF1 has a region strongly rich in A:T containing the hom-thrB promoter and site of action of the methionine mediated repression.

The DNA sequence of the chromosomal DNA insert in pFS80, encoding threonine synthase (thrC), was also determined by dideoxy sequencing techniques and is shown in Figure 4. A restriction map was predicted and checked against restriction analysis results to corroborate the accuracy of the sequence data. Computer aided analysis (UWGCG Programs, UW

Biotechnology Center, University of Wisconsin) was used to predict the thrC gene within the sequence data. These results were compared with in vivo genetic deletion analysis. An open reading frame extends 5' to GTG at 396) marking the amino terminal region of thrC.

The threonine synthase activity maps within the 1.57 kb Bcl1-Stu1 restriction fragment. The computer predicted structural gene sequence, GTG (396) to TAA (1881), lies completely within this fragment. The StuI restriction is 176 bp 3' of the preducted translation stop codon.

Heterospecific genetic complementation of the E. coli thrC 1001 auxotroph shows that the C. glutamicum thrC gene is expressed in E. coli. By comparison using computer searches for regions similar to E. coli ribosome binding sites and translation terminator sequences, a ribosome binding site adjacent to GTG(396) and a significant terminator-like sequence 35 bp 3' of the TAA at 1881 were identified. Homology was detected between C. glutamicum and E. coli thrC regions at both DNA and predicted protein sequence levels. Limited conservation of DNA sequence was observed between the E. coli thrC gene and the region 400 to 1400 bp of the C. glutamicum thrC sequence. There is consistent conservation in the central region (residues 100 to 350 of C. glutamicum thrC) and the carboxy terminal residues 430 to 480.

Identification of the hom-thrB transcription start site by S1 nuclease mapping and deletion analysis.

The transcriptional start site for the C. glutamicum hom-thrB genes was identified using S1 nuclease mapping, as described by Berk and Sharp, Cell 12, 721 (1977). The procedure requires the isolation and denaturation of a DNA fragment which overlaps the promoter and has been ^{32P} label at the 5'end of the antisense strand. Hybridization of this fragment to its cognate mRNA and subsequent digestion with the single strand specific exo-nuclease S1 results in the degradation of the 3'end of the labled DNA fragment up to the point at which it is protected by the RNA. The size of the result-ing DNA fragment is determined by comigration with DNA fragments resulting from the sequencing reac-tions of Maxam and Gilbert, Methods in Enzymol. 65, 499-559 (1982). This enables the identification of the transcriptional start site. The results can then be confirmed by deletion analysis of the promoter using restriction enzymes and exonuclease Bal 31 to construct series of deletions which are then reinserted into the organism and assayed for activity.

The Sma1-Hind111 restriction fragment that encompasses the hom-thrB promoter/operator and the first seven amino acid residues of the hom gene product was used in the S1 nuclease mapping studies. Plasmid pRA1 (pUC18 containing the 3.6 kb Sal1 C. glutamicum genomic fragment encoding hom-thrB) was

cut with Hind111 to generate a 1.014 kb restriction fragment, dephosphorylated with CIP (calf intestine phosphatase, Boehringer-Mannhein Biochemicals, Indianapolis, IN), labelled by treatment with polynucleotide kinase and gamma^{32} P-ATP (specific activity greater than 5000 Ci/mmol, Amersham Corp. Arlington Heights, IL), subsequently cleaved with Sma1 (New England Biolabs, Beverly, MA) to produce a 242 bp DNA fragment that was then purified by preparative polyacrylamide gel electropheresis. All manipulations were carried out in accordance with procedures described in Molecular Cloning by T. Maniatis et al. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982) and the enzyme suppliers recommendations.

Understanding the expression of a gene requires the isolation and structural characterization of the mRNA product, the size and number of transcripts, the regulatory control and the site of transcription initiation. Criteria evaluated for RNA isolated from Corynebacteria include RNA quality (Abs_{260}: Abs_{280}, ratio of about 1.95 to 2.05), purity (degradation and contamination determined by agarose gel electropheresis), and yield in mg RNA/liter cells (Abs_{260} = 4 micrograms RNA/ml).

RNA was extracted from C. glutamicum AS019 using the guanidinium isothiocyanate/French press isolation method. In this method, a 1 l. LB culture of AS019 is grown at 30°C to late exponential phase and harvested by 10 minutes centrifugation at 5000

RPM in a Sorvall GSA-250 rotor. The cells are washed at 4°C in 0.1 M NaCl, 10 mM Tris.Cl, pH 8.0, 1 mM EDTA, and harvested again. The pellets are combined in 50 ml 4 M guanidinium isothiocyanate, 2-mercaptoethanol (GuT:2ME) and immediately lysed by compression through a French press at approximately 1500 psi. Cell debris is sedimented by centrifugation in a Sorvall SS-34 rotor for 10 minutes at 10,000 rpm. Six ml aliquots of the supernatant are applied to 4 ml of 5.7 M CsCl, 10 mM ETTA, 25 mM sodium acetate and centrifuged at 34,000 rpm in a Beckman Ti50 fixed angle rotor for 24 hours.

The density gradient separates the sheared DNA molecules from the RNA, which forms a pellet at the tube base. This RNA pellet is resuspended in 5 ml 10 mM Tris, pH 7.5, 1 mM EDTA 5.0% Sarkosyl (TESK) containing 5.0% phenol. The solution is made 0.1 M with 5 M NaCl, and extracted with 10 ml 50% phenol, 49% chloroform, 1% isoamylalcohol (PCIA). The phases are separated by centrifugation in a Sorvall SS-34 at 3,000 rpm for 5 minutes and the phenolic phase back extracted with TESK containing 0.1 M NaCl. The combined aqueous phases are made 0.2 M with sodium acetate, pH 5.5, and the RNA precipitated overnight at -20°C in 2.5 volumes of ethanol. After centrifugation at 10,000 rpm for 20 minutes at 4°C, the RNA pellet is washed in ethanol, dried under vacuum, and resuspended in RNase free water at a concentration of 0.5 mg/ml.

Total cellular RNA, isolated from AS019 grown in minimal media with and without L-methionine (400 microgram/ml) supplementation, was separated by agarose gel electrophoresis, transferred to nitrocellulose paper and probed either with pMF-L2 or pUC-B5. Plasmid pMF-L2 contains a 1.8 kb Nae1 fragment which spans both the hom and thrB genes but contains no flanking sequences. RNA is glyoxylated to prevent spurious electrophoretic patterns caused by potential secondary structure. For each lane, 20 micrograms of C. glutamicum RNA is suspended in 8 microliters of glyoxal reaction mixture (1 M glyoxal, 50% DMSO, 10 mM potassium phosphate, pH 7.0) and incubated 1 hour at 50°C. Glyoxylated RNA samples are prepared for loading by the addition of 17 microliters formamide, 6.2 microliters formaldehyde, 3 microliters 10x running buffer 0.2 M morpholinopropanesulfonic acid (MPOS), 50 mM sodium acetate, 10 mM EDTA) and 5 microliters loading dye (50% glycerol, 1 mM EDTA, 0.4% bromophenol blue, 0.4% xylene cyanol). Samples are loaded onto an agarose/formaldehyde gel (2.2% agarose, 1 x running buffer, 18% formaldehyde, pH adjusted to 7.0 with NaOH) and electrophoresed at 30 mA. HindIII restricted lambda DNA is labeled with ^{32}P, denatured and glyoxylated similar to RNA samples and utilized as a molecular size standard. Following electrophoresis, nucleic acid is transferred to a nitrocellulose filter using the technique of Southern, J. Mol. Biol. 98, 503-517 (1975) except that no prior

treatment of the gel was necessary. Following transfer for 15 hours, the filters are baked in vacuo at 80°C for 2 hours.

The filters are prehybridized in sealed plastic bags for 16 hours at 42°C in a minimum volume, approximately 10 mls of hybridization buffer (50% deionized formamide, 5 x SSC, 50 mM sodium acetate, pH 6.5, 25 micrograms sonicated denatured salmon sperm DNA, 0.02% bovine serum albumin, 0.02% Ficoll, 0.02% polyvinyl pyrrolidone). (SCC = 0.15 M NaCl. 0.015 M sodium citrate, pH 7.0). The DNA probe is labeled with ^{32}P by nick translation, according to Rigby et al., J. Mol. Biol. 113, 237-251 (1977), heat denatured, added to the hybridization buffer, and incubated with the filter for 20 hour at 42°C. Filters are subsequently washed five times in 2 x SSC/0.1% SDS at room temperature and then three times in 0.2 SSC at 50°C. After drying, the filters are exposed to X-ray film and specific bands of hybridization determined by autoradiography.

Hybridization of pMF-L2 to total C. glutamicum RNA leads to the appearance of a single 2.4 kb transcript. This observation is in agreement with the predicted size of the hom-thrB transcript (2408 base pairs), based on S1 nuclease mapping, and the computer predicted termination site of the thrB gene. The size of the observed transcript and the lack of a detectable second transcript hybridizing to the hom-thrB probe leads the conclusion that C. glutamicum expresses hom and thrB from a single

transcriptional unit, representing the first defined operon in this organism. Results obtained by S1 nuclease mapping of the thrA-thrB junction support this conclusion.

Hybridization of hom-thrB specific RNA to the 242 bp Sma1-HindDIII end-labeled probe was achieved by lyophilizing 30 micrograms C. glutamicum RNA with 10 ng probe DNA and resuspending in 10 microliters hybridization buffer (40 mM PIPES pH 6.4, 0.4 M NaCl, 1 mM EDTA, 80% deionized formamide). The DNA was denatured by heating at 90°C for 10 minutes. Hybridization was performed overnight at 49°C.

The hybrid DNA-RNA molecules were digested with 2,000 units S1 nuclease (Bethesda Research Laboratories, Inc., Gaithersburg, MD) in 235 microliters assay buffer (250 mM NaCl, 30 mM sodium acetate, 10 mM zinc sulphate, 200 microgram/ml calf thymus DNA). The digest was incubated at 37°C for one hour and terminated by extraction with 250 microliters PCIA (phenal/chloroform/isoamyl alcohol, 50:48:2). Nucleic acids were precipitated from the aqueous phase with 0.2 M sodium acetate, 2 micrograms yeast tRNA and 50 microliters ethanol at -20°C. Following centrifugation and drying, each sample was dissolved in 3 microliters formamide loading buffer (100 ml formamide, 0.72 g Na_2 EDTA, 0.03 g bromophenol blue) and applied to a 6% polyacrylamide/7 M urea sequencing gel. The 242 bp Sma1-Hind111 restriction fragment was sequenced using the procedures for the

G, C+T and C reactions as described in Maxam and Gilbert, Methods in Enzymol. 65, 499-559 (1982).

The Sma1-Hind111 fragment labeled at the Hind111 5'terminus (antisense strand) acts as a specific probe to RNA complementary to this region. To detect the start of the hom-thrB mRNA transcript, total RNA is hybridized to the labeled probe and the unprotected single stranded nucleic acid digested with the single strand specific S1 nuclease. The length of the protected region of the DNA probe is resolved on a denaturing 6% polyacrylamide gel. The hom-thrB transcript initiates at coordinate 906, at the first of a GG doublet. This nucleotide is 88 bp 5' of the first available ATG codon in the hom open reading frame. This defines the promoter region responsible for hom-thrB expression and is designated thrP1 for threonine promoter 1. The sequence is shown in Figure 5a.

No detectable degradation of the DNA probe from the 0.46 kb Fok1- Pvu11 restriction fragments spanning the hom thrB Pvu11 junction, indicating that the majority of the thrB expression was mediated by thrP1.

Identification and deletion of the operator mediating L-methionine repression of hom-thrB expression.

In addition to promoter identification, restriction and/or exonuclease Bal31 deletions have been utilized in identification and deletion of the operator (thrO), which mediates the L-methionine

repression of hom-thrB expression and in the construction of a feedback inhibition deficient variant of the hom gene. These studies were facilitated by construction of a special vector designated pWST1. When investigating promoter structure and function on a plasmid, it is desirable to eliminate read through transcription from upstream promoters located within the cloning vector. Plasmid pWST1 contains the E. coli trpA terminator followed by a polylinker to facilitate the cloning of the gene in the various deletion generated variants. The effect of the deletions can be assayed in the absence of influence by upstream promoters. This vector is applicable not only to the analysis of the hom-thrB genes, but also the characterizations of other promoter/operator systems in C. glutamicum. pWST1 is constructed using the trpA terminator obtained from Pharmacia Fine Chemicals, Piscataway, NJ. Sac1 linkers are attached and the trpA terminator inserted into the polylinker region of M13mp19. The constructs are sequenced using the method of Sanger et al. Proc. Natl. Acad. Sci. USA 74,5463-5467 (1977), to screen for insertion of the terminator in the proper orientation. The terminator/polylinker is subsequently ligated into the Sma1-Sal1 restricted pTF33, a derivative of the C. glutamicum/E. coli shuttle vector pWS124, described by Batt et al., Biotechnol. Letts. 7:717 (1985). DNA linkers and enzymes are obtained from New England Biolabs or Boehringer-Mannheim, as noted earlier.

The hom-thrB genes ligated into Sma1/Sal1 restricted pWST1 on a 2911 bp Sma1-Sal1 restriction fragment and a 2,815 bp Dra1-Sal1 restriction fragment, designated pWFS2.9 and pWFS2.8, respectively. Further deletion of the hom-thrB upstream region is accomplished as diagramed in Figure 6. The recombinant vector pWFS2.9 was linearized by Sma1 digestion and deletions constructed by digestion of 6 micrograms of DNA with the exonuclease Bal31 (0.2 units)/ micrograms DNA. Aliquots of the reaction mixture were removed at 30 second intervals between 4 and 15 minutes, and the reaction stopped by dilution into one volume of 50 mM EDTA. The DNA was digested with Sal1 and the resulting hom-thrB containing fragments purified by agarose gel electrophoresis. These fragments were ligated into Sma1-Sal1 digested pWST1 and the resulting recombinant mixture used to transform C. glutamicum AS253 (hom).

The extent of the Bal31 generated deletions, diagrammed in Fig. 7, in complementing a non-complementing derivative plasmid is determined by nucleotide sequence analysis and measurement of the levels of homoserine dehydrogenase activity in crude extracts. The ability of the deletion plasmids to complement the hom-thrB auxotrophy of strain AS253 was checked by streaking the corresponding AS253 transformants onto MCG/kanamycin agar plates.

The results of the deletion construction and their effect on hom gene expression show that deletion of sequences upstream of hom, up to the

Sma1 (pWFS2.9) or Dra1 (pWFS2.8) restriction site (218 and 124 bp prior to the predicted translational start site, respectively) does not drastically influence the expression of the hom gene.

As predicted from the S1 nuclease studies, the Dra1-Sal1 hom-thrB fragment contains both the -10 and -35 regions which are critical for promoter activity in E. coli, as reported by Hawley and McClure, Nucleic Acid Res. 11:2232 (1983). Further deletion by Bal31 markedly reduces the expression of the hom gene product homoserine dehydrogenase, supporting the data obtained with S1 nuclease mapping of thrP1. Two of the deletion derivatives, pWFSdelta2304 and pWFSdelta2207 are able to complement C. glutamicum AS253 despite the reduction in hom expression to 2 and 10%, respectively, of that observed in strains containing the parental plasmid pFS3.6. The relative specific activity of homoserine dehydrogenase observed in C. glutamicum AS253 (hom) harboring these two deletion derivatives is 3.1 and 0.7 with respect to that observed in wild type strains. C. glutamicum auxotrophs requiring threonine/methionine express approximately 2% of the wild type level of homoserine dehydrogenase activity. The deletion of the hom promoter carried in pWFSdelta2431 results in a 96-fold decrease in the expression of the cloned thrB gene thus demonstrating a common promoter.

The Bal31 generated deletions enable the mapping of the boundary between those deletion

derivatives which complement the threonine/ methionine auxotrophy and those that fail to complement, between 63 base pairs (pWFSdelta2304) and 56 base pairs (pWFSdelta2431) upstream of the predicted hom translation start sites. The observation that deletions extending to locations beyond the putative start point of transcription (88 bp upstream of the hom start codon) is determined by S1 nuclease mapping, does not necessarily result in the complete loss of homoserine dehydrogenase activity. This loss may also be due to weak promoter activity adjacent to the main transcription start site.

The mechanism of transcriptional regulation of the hom-thrB operon was determined to involve control by the single stem/loop attenuator shown in Figure 5. Specific deletion of the stem/loop structure removes the methionine repression of hom-thrB expression. In this structure, the sequence ATGTAG, encoding Met-Stop, forms the loop. The sequence TTTTGGACA, similar to the TTGGAGA that precedes the predicted translational start site of the hom gene, precedes the ATG and thus represents a potential ribosome binding site. A possible model is one in which a bound ribosome can be momentarily stalled due to a low concentration of charged methionine tRNA, thus preventing stem/loop formation and allowing transcription to continue. At higher concentrations of methionine, the ribosome would move to the TAG Stop signal and disengage, allowing stem/loop formation.

LacZ protein fusions can be used to evaluate promoter and operator functions directly in C. glutamicum. This method is demonstrated using the hom-thrB promoter/operator, isolated on a 330 bP Sma1-Hind111 fragment purified and ligated into similarly digested pSKS107 that contains a promoterless lactose operon. The construction creates a hom-lacZ protein fusion containing the N-terminal eight hom amino acid residues preceding the lacZ gene product beta-galactosidase. The expression of the fusion protein products required the insertion of a ribosome binding site, initiating codon (ATG/GTG), under the control of the hom-thrB promoter/operator. The recombinant vector was introduced into E. coli JM83 where beta-galactosidase activity as observed in crude extracts. Supplementation of the growth medium with L-methionine represses the expression of lacZ two-fold.

Deletion of a portion of the dyad symmetry element required for operator function demonstrates the role of the dyad symmetry element in the regulation of the hom-thrB gene. A 1.48 kb Kpn1 restriction fragment containing the hom gene of the Bal31 deletion derivative, pWFSdelta2207 was purified and used to replace its counterpart in the parental vector pFS3.6A. The resulting recombinant plasmid, designated pWFS2207delta1, contains a specific 10 base pair deletion removing the left half of the dyad symmetry element. Identical levels

of homoserine dehydrogenase activity were measured in strains grown in MCG medium with and without methionine supplementation. This demonstrates that the dyad symmetry element is the site of L-methionine repression (thrO).

Identification of the thrC promoter by deletion analysis, complementation of auxotrophs and overproduction of the enzyme.

The promoter sequence for expression of thrC was also determined by deletion analysis, auxotrophic complementation, and overproduction of enzyme.

The overproduction of the product of the thrC gene, threonine synthetase, can be measured from crude extracts of C. glutamicum strains AS213 and wild type AS019 containing the parental vector pHY416 or the thrC containing pFS80. The two strains containing the plasmids are grown in MCG medium, the cells harvested, lysed, cell debris removed by centrifugation, the protein purified by 40 to 60% ammonium sulphate fractionation, DEAE-Sephadex column chromatography with a 0.2 M to 0.2M to 0.6 M KCl gradient, and anion exchange chromatography in a FPLC column dluted with a 0.1 M to 0.7 M KCl gradient. The results demonstrate that the protein is produced at a level 200% of that observed in the wild type. The method produces threonine synthetase specific activity demonstrating a purification of over 350-fold. The protein has a

molecular weight of 56,000 by SDS polyacrylamide gel electropheresis.

Analysis of activity complementation of the thrC auxotroph indicates that the promoter for the thrC gene sequence precedes the predicted translation start site for the thrC gene product by approximately 80 base pairs. The sequence, TTGAAA--(16 bp)--TAGGGT, is closely related to the E. coli consensus sequence as well as the promoter sequence determined for C. glutamicum thrP1, AAAGCA--18bp--TATAGT. Confirmation of the identification of the sequences the thrC gene promoter sequence is done by S1 nuclease analysis.

Modification of the enzyme structure and expression of hom, thrB and thrC.

Once the hom, thrB and thrC genes are identified, including the identification if the initiation sites of both mRNA and protein synthesis for the genes, it is possible to increase the quantity of gene expression by increasing the gene dosage by localization of specific genes on a multicopy plasmid, by site-directed mutagenesis or replacement of the promoter, by increasing translational efficiency through alteration of the ribosome binding site, or by increasing stability of the protein by site-directed mutagenesis. The quality of the particular gene can be increased using in vitro and site-directed mutagenesis to alter substrate utilization as well as the kinetic and regulatory properties of the enzyme. Physical properties such

as heat stability can also be modified. The construction of a vector with the three genes transcribed as a single unit under the control of a high efficiency promoter results in more efficient threonine synthesis. One can also remove the L-threonine feedback inhibition of homoserine dehydrogenase or the threonine and lysine inhibition of aspartokinase to produce the overproduction of threonine. The feedback inhibition of homoserine dehydrogenase can be removed by in vitro mutagenesis using either hydroxylamine and/or sodium bisulfite, methods well known to those skilled in the art, or by recombinant techniques. The mutagenized plasmids are reintroduced into C. glutamicum and screened for AHV resistance or by enzyme assays. Increased promoter efficiency can also be accomplished by site directed mutagenesis of the existing promoter or by replacement with a high efficiency promoter.

The thrC gene can be placed under the transcriptional control of high efficiency promoters such as the E. coli promoter tac to produce elevated levels of the gene product. The expression vector pKK233-2, obtained from Pharmacia Fine Chemicals, is restricted with Nco1-Hind111. The plasmid PFS80 is cleaved with Bcl1, blunt ended with Klenow polymerase, and Nco1 linkers ligated onto the flush ends. The ligation product is double digested with Nco1 and Hind111, a 2.8 kb fragment purified and ligated into similarly digested pKK233-2. The resulting recombinant vector designated pKC14 is

transformed into E. coli JM105. The requirement for IPTG (isopropyl-beta-D-thio-galactopyranoside) induction demonstrates that the thrC gene is under the transcriptional control of the tac promoter. The threonine synthetase activity measured in the absence of IPTG was 1.3 nmole/min/mg protein. The addition of 2 mM IPTG induces 24 times the threonine synthetase (30.9 nmole/min/mg-protein).

The homoserine dehydrogenase and homoserine kinase polypeptides encoded by the open reading frames corresponding to the hom and thrB gene products can be expressed and purified for analysis. The enzymes are purified from 10 liters of MCG broth from a CHEMAP fermentator innoculated with a 300 ml overnight culture of C. glutamicum AS019/pFS3.6 grown for 24 hours at 30°C with 470 rpm agitation. Cells were harvested by ultrafiltration and centrifugation, the cell pellet resuspended in lysis buffer (100 mM KPO_4, pH 7.0, 0.5 M KCl) and the cells lysed by repeated passage through a French pressure cell. Debris is removed by centrifugation, the supernatant precipitated with ammonium sulphate, and the enzyme activities separated on a DEAE-Sephadex A-50 column eluted with a linear 0.3 M to 0.8 M KCl gradient. Fractions containing the appropriate enzyme activity are pooled and the proteins analyzed by SDS polyacrylamide electrophoresis. The proteins are then further purified on hydroxylapatite HPLC prior to final separation by preprative SDS-PAGE. The purified homoserine

dehydrogenase remained active through the procedure and has a final molecular weight of 47,000 daltons. The activity of the homoserine kinase is lost, but the protein had a molecular weight of 32,000.

The observed molecular weights are in close agreement with the molecular weights predicted from the nucleotide sequences. The NH_2 terminus of the homoserine dehydrogenase is blocked, however, the amino acid composition is in good agreement with the amino acid composition predicted from the gene sequence. The expression of activity by ligation of the SmaI-HindIII fragment containing the C. glutamicum thrP1 and predicted NH_2-terminal seven amino acid residues of the hom gene product indicate that the N-terminal sequence is correct. The first ten residues of the thrB gene product, homoserine kinase, is in complete agreement with the predicted amino acid sequence. This identifies the translation initiation site for the thrB gene as the ATG at nucleotide 2342, confirming the predicted primary structure of the C. glutamicum homoserine kinase. The protein appears to undergo post-translational removal of the N-formyl-MET, a relatively common feature of procaryotic proteins.

Construction and application of a C. glutamicum promoter probe.

A promoter probe for use in identifying, isolating, and quantifying the efficiency of promoters is based on pWST1 and designated pAL-1. The chloramphenicol acetyltransferase (CAT) gene is

used as the promoter probe gene. The cat gene is expressed in E. coli., B.subtilis, and C. glutamicum. Since cloning of strong promoters can induce plasmid instability by transcriptional interference at the replication origin, it may be necessary to clone the fd enteric bacteriophage major gene terminator at the 3' terminus of the test gene. Potential promoter sequences can then be screened for the acquisition of chloramphenicol resistance. Preliminary estimation of promoter efficiency can be accomplished by determining the extent of antibiotic resistance.

A number of promoter sources can be screened for their efficiency in C. glutamicum. High efficiency promoters from other procaryotic systems are known, for example, E. coli trp, the hybrid Ptac, lambda pH, and B. subtilis Preg. Random C. glutamicum chromosomal DNA and/or corynephage DNA fragments can also be inserted into the polylinker site, upstream in the test gene and cat activity determined to assess promoter efficiency. DNA sequencing and determination of the transcriptional start sites are used to characterize the promoter structure. For example, the 266 bp DraI-Hae111 fragment which spans the predicted thrC promoter region was purified and ligated into the Sma1 restriction site of the pAL-1 polylinker. The resulting recombinant mixture was introduced into E. coli and chloramphenicol resistant transformance obtained.

The primary site of metabolic regulation of threonine biosynthesis in C. glutamicum is the L-threonine inhibition of homoserine dehydrogenase. To remove the metabolic block, in vitro mutogenesis has been used to alter the hom gene product to produce feedback inhibition deficient variants. The recombinant plasmid pWFSdelta2207 was used as a source of the hom gene. This plasmid expresses a lower level of homoserine dehydrogenase (10%) than pFS3.6A, eliminating potential artifacts due to the overproduction of the hom gene product. Plasmid DNA was digested with Kpn1, separated by agarose gel electrophoresis and the fragments isolated by electroelution into dialysis bags. The 1.43 kb Kpn1 fragment containing the hom gene was purified and treated with hydroxylamine. This is a potent mutagen primarily causing AT to GC and GC to AT transitions. The mutated hom gene was isolated and 3 micrograms of target DNA resuspended in 280 microliters of 1 M hydroxylamine, 0.3 M KPO_4, pH 6.0, aliquots removed at between 10 and 300 minutes and the reaction stopped by ethanol precipitation. The mutagenized fragments were religated to the large, 12.7 kb, Kpn1 restriction fragment and transformed into E. coli JM83.

Plasmid DNA from ampicillin resistant transformants was purified on CsCl gradients and transformed into the restriction deficient C. glutamicum AS019-E12. These transformants were then screened for resistance to alpha-aminohydroxyvaleric acid

(AHV) in order to select for a deregulated hom gene product. Homoserine dehydrogenase activity assays were used to confirm and demonstrate the removal of the L-threonine mediated feedback inhibition. Different mutations could be combined by recombinant DNA techniques to determine the extent to which they are cooperative.

The present invention, nucleotide sequences encoding threonine biosynthetic enzymes, and methods and sequences for the expression and regulation of expression of these enzyme encoding sequences are disclosed. Modifications and variations of this invention will be obvious to those skilled in the art of genetic engineering from the foregoing detailed description. It is intended that these modifications and variations will fall within the scopes of the appended claims.

WE CLAIM:

1. A method for the production of threonine comprising:

providing nucleotide sequences for the genes encoding the enzymes in the threonine biosynthetic pathway; and

inserting the enzyme encoding sequences with selected nucleotide sequences mediating the expression and regulation of the enzyme encoding sequences into an expression vector.

2. The method of claim 1 further comprising inserting said expression vector into an expression host.

3. The method of claim 2 wherein said expression host is Corynebacterium.

4. The method of claim 1 wherein said nucleotide expression and regulation sequences include a promoter, further comprising selecting a promoter having a higher efficiency than the promoter associated with the enzyme encoding chromosomal genes and selecting for greater efficiency.

5. The method of claim 4 wherein said higher efficiency promoter is obtained by mutating the promoter associated with said enzyme encoding genes.

6. The method of claim 4 further comprising selecting said high efficiency promoter from the group of promoters isolated from Escherchia, Bacillus, Staphylococcus and Streptococcus.

7. The method of claim 1 further comprising selecting a multicopy plasmid as the expression vector.

8. The method of claim 1 wherein said expression and regulation sequences include a ribosome binding site, further comprising selecting for a ribosome binding site with increased efficiency.

9. The method of claim 1 further comprising combining said enzyme encoding nucleotide sequences in a single expression vector.

10. The method of claim 1 further comprising mutating said enzyme encoding nucleotide sequences and selecting for temperature stability.

11. The method of claim 1 further comprising mutating said enzyme encoding nucleotide sequences and selecting for substrate utilization.

12. The method of claim 1 wherein said expression and regulation signals include a repressor, further comprising modifying said repressor.

13. The method of claim 12 wherein said repressor is deleted.

14. The method of claim 12 wherein said repressor is mutated.

15. The method of claim 12 wherein said repressor is replaced with a repressor other than the repressor associated with the chromosomal gene encoding said sequence.

16. A method of constructing a promoter probe for Corynebacterium amino acid genes comprising:
isolating nucleotide sequences encoding a detectable protein product involved in threonine biosynthesis in a Corynebacterium host,
constructing deletions of the 3' end of said nucleotide sequences,
inserting said deletions into an expression vector,
transforming said vector into an auxotrophic Corynebacterium host, and
determining if said deletions produce protein in auxotrophic Corynebacterium hosts.

17. The method of claim 16 further comprising sequencing the nucleotide sequences complementing the auxotrophic Corynebacterium hosts.

18. A rDNA sequence comprising hom, thrB, and thrC.

19. The rDNA sequence of claim 18 further comprising a promoter sequence.

20. The rDNA sequence of claim 19 further comprising a repressor sequence.

21. The rDNA sequence of claim 19 wherein said promoter sequence is selected from sequences having a higher translational efficiency than the sequences associated with the chromosomal DNA.

22. A Corynebacterium rDNA promoter sequence comprising a ribosome binding site TTGGAGA.

23. A nucleotide sequence hybridizing to a rDNA sequence encoding homoserine dehydrogenase in Corynebacteria.

24. A nucleotide sequence hybridizing to a rDNA sequence encoding homoserine kinase in Corynebacteria.

25. A nucleotide sequence hybridizing to a rDNA sequence encoding threonine synthase.

26. A Corynebacterium rDNA translation termination sequence comprising
A A G G A A G G C C C C T T C G A A T C A A G A
A G G G G C C T T.

27. A Corynebacterium rDNA translation termination sequence comprising
G A T G G A A C C A G G C C T T T C G C A T T G
A G T G G C G T T T T A A G G C C T C C A.

28. A Corynebacterium rDNA sequence repressing translation in the presence of excess methionine comprising
T T T G T T T T G G A C A C A T G T T C T A G G
met stop
G T G G C C G A A A C A A A.

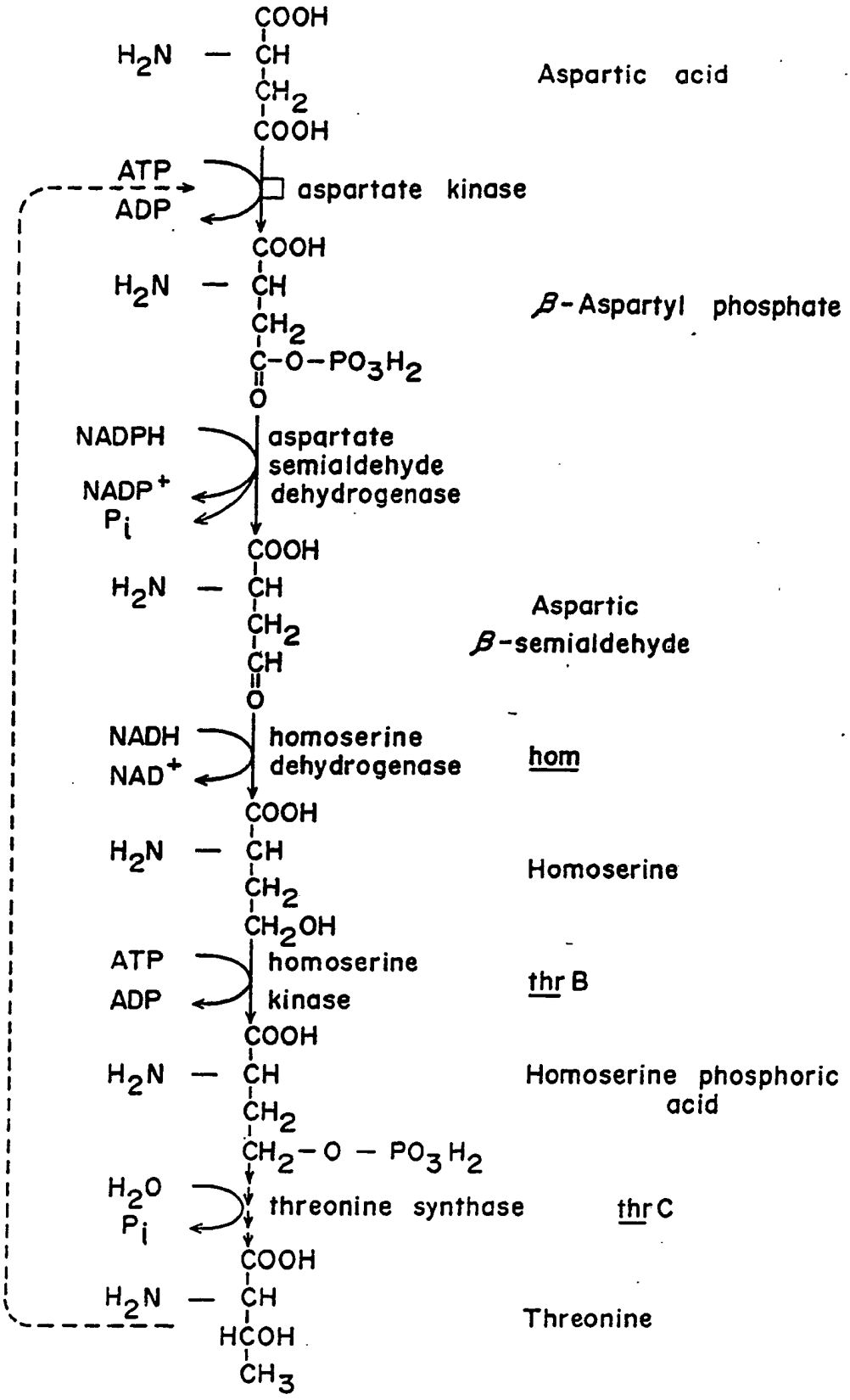

Aspartic acid
ATP
ADP
aspartate kinase
β-Aspartyl phosphate
NADPH
NADP+
Pi
aspartate semialdehyde dehydrogenase
Aspartic β-semialdehyde
NADH
NAD+
homoserine dehydrogenase
hom
Homoserine
ATP
ADP
homoserine kinase
thr B
Homoserine phosphoric acid
H2O
Pi
threonine synthase
thr C
Threonine

FIGURE 1

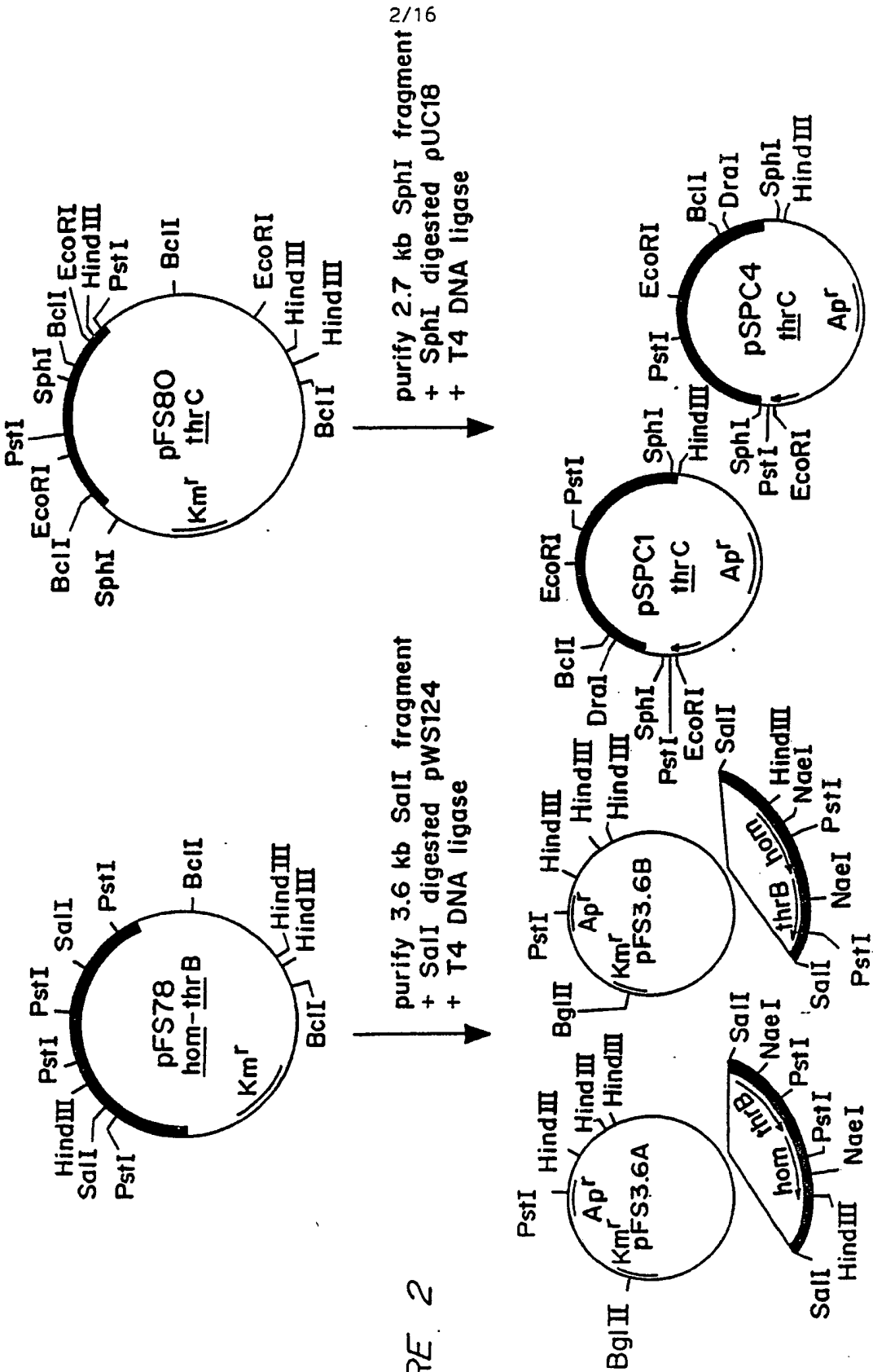

pFS80
thrC
Km^r
purify 2.7 kb SphI fragment
+ SphI digested pUC18
+ T4 DNA ligase
pFS78
hom-thrB
Km^r
purify 3.6 kb SalI fragment
+ SalI digested pWS124
+ T4 DNA ligase
pSPC1
thrC
Ap^r
pSPC4
thrC
Ap^r
pFS3.68
Ap^r
Km^r
pFS3.6A
Ap^r
Km^r
hom
thrB

FIGURE 2

FIGURE 3

```
   1  GTCGACCGCGTGAAGTCGCCCTTTAGGAGAATTCTGACTAACTGGAGCCAAAACTTGATC   60

  61  CACTCGAGAGCTGTGCAGTCTCTTTTTCCTTCAATTCTGCCTGCTCGAGCTCGTAGAAGT  120

 121  AGAGGTCTACTTCAGTTGGTTCACCTTGCACACAAGCATGAAGTAGTGGGTAGGTCGAGT  180

 181  TGTTAAATGCGGTGTAGAAGGGGAGTAGTTCGCTAGCAAAGGTTAATTTGGAGTCGCTGT  240

 241  ACTGCGGGTTCTCGGGTGGAGTATTCCCGGAGGATTCAAGAAATCTTGACGCATCTTTGA  300

 301  TGAGGTATGTTTGGAATTCGTCGGCACCTTCCTCGCCGGAGAGGTAGTAGGAGTTGTCGT  360

 361  AATTTGGAACCCAGATGGCAAATCGTGCGTTTTCGATTGCGTCCAGGACTTCCTCTACGT  420

 421  TGTATCTCGCACTTGTTGCAGCGGAAGCGACTCGGTTGCCGATGTCTCCGTATGCAGTGA  480

 481  GCGTGGCGTTTCCGAGGGGAACTTGATCAGAGGAATACACCATGGAGCCGATGTCAGAGG  540

 541  CGACTGCGGGCAGATCCTTTTGAAGCTGTTTCACAATTTCTTTGCCCAGTTCGCGGCGGA  600

 601  TCTGGAACCACTTTTGCATGCGATCGTCGTCAGAGTGGTTCATGTGAAAAATACACTCAC  660

 661  CATCTCAATGGTCATGGTGAAGGCCTGTACTGGCTGCGACAGCATGGAACTCAGTGCAAT  720

 721  GGCTGTAAGGCCTGCACCAACAATGATTGAGCGAAGCTCCAAAATGTCCTCCCCGGGTTG  780
```

FIGURE 3 (CONT'D)

```
781  ATATTAGATTTCATAAATATACTAAAAATCTTGAGAGTTTTTCCGTTGAAAACTAAAAAG  840

841  CTGGGAAGGTGAATCGAATTTCGGGGCTTTAAAGCAAAAATGAACAGCTTGGTCTATAGT  900

901  GGCTAGGTACCCTTTTTGTTTTGCACACATGTAGGGTGGCCGAAACAAAGTAATAFFACA  961

                                          MetThrSerAlaSerAlaProSerPhe
961  ACAACGCTCGACCGCGATTATTTTTGGAGAATCATGACCTCAGCATCTGCCCCAAGCTTT 1020
                                          Translation Initiation Codon

     AsnProGlyLysGlyProGlySerAlaValGlyIleAlaLeuLeuGlyPheGlyThrVal
1021 AACCCCGGCAAGGGTCCCGGCTCAGCAGTCGGAATTGCCCTTTTAGGATTCGGAACAGTC 1080

     GlyThrGluValMetArgLeuMetThrGluTyrGlyAspGluLeuAlaHisArgIleGly
1081 GGCACTGAGGTGATGCGTCTGATGACCGAGTACGGTGATGAACTTGCGCACCGCATTGGT 1140

     GlyProLeuGluValArgGlyIleAlaValSerAspIleSerLysProArgGluGlyVal
1141 GGCCCACTGGAGGTTCGTGGCATTGCTGTTTCTGATATCTCAAAGCCACGTGAAGGCGTT 1200

     AlaProGluLeuLeuThrGluAspAlaPheAlaLeuIleGluArgGluAspValAspIle
1201 GCACCTGAGCTGCTCACTGAGGACGCTTTTGCACTCATCGAGCGCGAGGATGTTGACATC 1260

     ValValGluValIleGlyGlyIleGluTyrProArgGluValValLeuAlaAlaLeuLys
1261 GTCGTTGAGGTTATCGGCGGCATTGAGTACCCACGTGAGGTAGTTCTCGCAGCTCTGAAG 1320

     AlaGlyLysSerValValThrAlaAsnLysAlaLeuValAlaAlaHisSerAlaGluLeu
1321 GCCGGCAAGTCTGTTGTTACCGCCAATAAGGCTCTTGTTGCAGCTCACTCTGCTGAGCTT 1380

     AlaAspAlaAlaGluAlaAlaAsnValAspLeuTyrPheGluAlaAlaValAlaGlyAla
1381 GCTGATGCAGCGGAAGCCGCAAACGTTGACCTGTACTTCGAGGCTGCTGTTGCAGGCGCA 1440
```

FIGURE 3 (CONT'D)

```
      IleProValValGlyProLeuArgArgSerLeuAlaGlyAspGlnIleGlnSerValMet
1441  ATTCCAGTGGTTGGCCCACTGCGTCGCTCCCTGGCTGGCGATCAGATCCAGTCTGTGATG 1500

      GlyIleValAsnGlyThrThrAsnPheIleLeuAspAlaMetAspSerThrGlyAlaAsp
1501  GGCATCGTTAACGGCACCACCAACTTCATCTTGGACGCCATGGATTCCACCGGCGCTGAC 1560

      TyrAlaAspSerLeuAlaGluAlaThrArgLeuGlyTyrAlaGluAlaAspProThrAla
1561  TATGCAGATTCTTTGGCTGAGGCAACTCGTTTGGGTTACGCCGAAGCTGATCCAACTGCA 1620

      AspValGluGlyHisAspAlaAlaSerLysAlaAlaIleLeuAlaSerIieAlaPleHis
1621  GACGTCGAAGGCCATGACGCCGCATCCAAGGCTGCAATTTTGGCATCCATCGCTTTCCAC 1680

      ThrArgValThrAlaAspAspValTyrCysGluGlyIleSerAsnIleSerAlaAlaAsp
1681  ACCCGTGTTACCGCGGATGATGTGTACTGCGAAGGTATCAGCAACATCAGCGCTGCCGAC 1740

      IleGluAlaAlaGlnGlnAiaGlyHisThrIleLysLeuLeuAlaIleCysGluLysPhe
1741  ATTGAGGCAGCACAGCAGGCAGGCCACACCATCAAGTTGTTGGCCATCTGTGAGAAGTTC 1800

      ThrAsnLysGluGlyLysSerAlaIleSerAlaArgValHisProThrLeuLeuProVal
1801  ACCAACAAGGAAGGAAAGTCGGCTATTTCTGCTCGCGTGCACCCGACTCTATTACCTGTG 1860

      SerHisProLeuAlaSerValAsnLysSerPheAsnAlaIlePheValGluAlaGluAla
1861  TCCCACCCACTGGCGTCGGTAAACAAGTCCTTTAATGCAATCTTTGTTGAAGCAGAAGCA 1920

      AlaGlyArgLeuMetPheTyrGlyAsnGlyAlaGlyGlyAlaProThrAiaSerAlaVai
1921  GCTGGTCGCCTGATGTTCTACGGAAACGGTGCAGGTGGCGCGCCAACCGCGTCTGCTGTC 1980

      LeuGlyAspValValGlyAlaAlaArgAsnLysValHisGlyGlyArgAlaProGlyGlu
1981  CTTGGCGACGTCGTTGGTGCCGCACGAAACAAGGTGCACGGTGGCCCTGCTCCAGGTGAG 2040
```

FIGURE 3 (CONT'D)

```
      SerThrTyrAlaAsnLeuProIleAlaAspPheGlyGluThrThrThrArgTyrHisLeu
2041  TCCACCTACGCTAACCTGCCGATCGCTGATTTCG,TGAGACCACCACTCGTTACCACCTC 2100

      AspMetAspValGluAspArgValGlyValLeuAlaGluLeuAlaSerLeuPheSerGlu
2101  GACATGGATGTGGAAGATCGCGTGGGGGTTTTGGCTGAATTGGCTAGCCTGTTCTCTGAG 2160

      GlnGlyIleSerLeuArgThrIleArgGlnGluGluArgAspAspAspAlaArgLeuIle
2161  CAAGGAATCTCCCTGCGTACAATCCGACAGGAAGAGCGCGATGATGATGCACGTCTGATC 2220

      ValValThrHisSerAlaLeuGluSerAspLeuSerArgThrValGluLeuLeuLysAla
2221  GTGGTCACCCACTCTGCGCTGGAATCTGATCTTTCCCGCACCGTTGAACTGCTGAAGGCT 2280

      LysProValValLysAlaIleAsnSerValIleArgLeuGluArgAsp
2281  AAGCCTGTTGTTAAGGCAATCAACAGTGTGATCCGCCTCGAAAGGGACTAATTTTACTGA 2340
                                                      Stop

      Predicated start of thrB translation
      MetAlaIleGluLeuAsnValGlyArgLysValThrValThrValProGlySerSerAl
2341  CATGGCAATTGAACTGAACGTCGGTCGTAAGGTTACCGTCACGGTACCTGGATCTTCTGC 2400
      Translation initiation Codon

      aAsnLeuGlyProGlyPheAspThrLeuGlyLeuAlaLeuSerValTyrAspThrValGl
2401  AAACCTCGGACCTGGCTTTGACACTTTAGGTTTGGCACTGTCGGTATACGACACTGTCGA 2460

      uValGluIleIleProSerGlyLeuGluValGluValPheGlyGluGlyGlnGlyGluVa
2461  AGTGGAAATTATTCCATCTGGCTTGGAAGTGGAAGTTTTTGGCGAAGGCCAAGGCGAAGT 2520

      LProLeuAspGlySerHisLeuValValLysAlaIleArgAlaGlyLeuLysAlaAlaAs
2521  CCCTCTTGATGGCTCCCACCTGGTGGTTAAAGCTATTCGTGCTGGCCTGAAGGCAGCTGA 2580

      pAlaGluValProGlyLeuArgValValCysHisAsnAsnIleProGlnSerArgGlyLe
2581  CGCTGAAGTTCCTGGATTGCGAGTGGTGTGCCACAACAACATTCCGCAGTCTCGTGGTCT 2640
```

FIGURE 3 (CONT'D)

```
      UGlySerSerAlaAlaAlaAlaValAlaGlyValAlaAlaAlaAsnGlyLeuAlaAspPh
2641  TGGCTCCTCTGCTGCAGCGGCGGTTGCTGGTGTTGCTGCAGCTAATGGTTTGGCGGATTT 2700

      EProLeuThrGlnGluGlnIleValGlnLeuSerSerAlaPheGluGlyHisProAspAs
2701  CCCGCTGACTCAAGAGCAGATTGTTCAGTTGTCCTCTGCCTTTGAAGGCCACCCAGATAA 2760

      nAlaAlaAlaSerValLeuGlyGlyAlaValValSerTrpThrAsnLeuSerIleAspGl
2761  TGCTGCGGCTTCTGTGCTGGGTGGAGCAGTGGTGTCGTGGACAAATCTGTCTATCGACGG 2820

      yLysSerGlnProGlnTyrAlaAlaValProLeuGluValGlnAspAsnIleArgAlaTh
2821  CAAGAGCCAGCCACAGTATGCTGCTGTACCACTTGAGGTGCAGGACAATATTCGTGCGAC 2880

      rAlaLeuValProAsnPheHisAlaSerThrGluAlaValArgArgValLeuProThrGi
2881  TGCGCTGGTTCCTAATTTCCACGCATCCACCGAAGCTGTGCGCCGAGTCCTTCCCACTGA 2940

      uValThrHisIleAspAlaArgPheAsnValSerArgValAlaValMetIleValAlaLe
2941  AGTCACTCACATCGATGCGCGATTTAACGTGTCCCGCGTTGCAGTGATGATCGTTGCGTT 3000

      uGlnGlnArgProAspLeuLeuTrpGluGlyThrArgAspArgLeuHisGlnProTyrAr
3001  GCAGCAGCGTCCTGATTTGCTGTGGGAGGGTACTCGTGACCGTCTGCACCAGCCTTATCG 3060

      gAlaGluValLeuProIleThrSerGluTrpValAsnArgLeuArgAsnArgGlyTyrAl
3061  TGCAGAAGTGTTGCCTATTACCTCTGAGTGGGTAAACCGCCTGCGCAACCGTGGCTACGC 3120

      aAlaTyrLeuSerGlyAlaGlyProThrAlaMetValLeuSerThrGluProIleProAs
3121  GGCATACCTTTCCGGTGCCGGCCCAACCGCCATGGTGCTGTCCACTGAGCCAATTCCAGA 3180

      pLysValLeuGluAspAlaArgGluSerGlyIleLysValLeuGluLeuGluValAlaGl
3181  CAAGGTTTTGGAAGATGCTCGTGAGTCTGGCATTAAGGTGCTTGAGCTTGAGGTTGCGGC 3240
```

FIGURE 3 (CONT'D)

YProValLysValGluValAsnGlnPro
3241 ACCAGTCAAGGTTGAAGTTAACCAACCTTAGGCCCAACAAGGAAGGCCCCTTCGAATCAA 3300

Stop Computer predicted

3301 GAAGGGGCCTTATTAGTGCAGCAATTATTCGCTGAACACGTGAACCTTACAGGTGCCCGG 3360
translation intermination point

3361 CGCGTTGAGTGGTTTGAGTTCCAGCTGGATGCGGTTGTTTTCACCGAGGCTTTCTTGGAT 3420

3421 GAATCCGGCGTGGATGGCGCAGACGAAGGCTGATGGGCGTTTGTCGTTGACCACAAATGG 3480

3481 GCAGCTGTGTAGAGCGAGGGAGTTTGCTTCTTCGGTTTCGGTGGGGTCAAAGCCCATTTC 3540

3541 GCGGAGGCGGTTAATGAGCGGGGAGAGGGCTTCGTCGAGTTCTTCGGCTTCGGCGTGGTT 3600

3601 AATGCCCATGACGTGTGCCCACTGGGTTCCGATGGAAAGTGCTTTGGCGCGGAGGTCGGG 3660

3661 GTTGTTGCATTGCGTCATCGTCGAC 3685

FIGURE 4

```
   1  TGACCGAGAGTTTTTTTGAGCAACTGGATCATTAGATAATTGTTCGATCGACCGAATGAA  60

  61  ATCACCCGTTATGGAGACCTACTGGAATTGAGCCCAGAAACCGTCGATGTGTGCCTCAAC 120

 121  GTAGGGGTAAAGCCACGGCCCGAGCAGCACCAGCCCGACCGCGAGCACCGAACAACCAAT 180

 181  GAGAACATACAGGTTCCACTTGGACACCGGCGCTGGATTAAGGATTTCAACTGCGGTGAG 240

 241  ATTCTTCTTGTTGTTGTCCTCGAGTTTCGAGAAGCTGGGGTAATCGGGAGCTGTCATCTT 300

 301  TAAAGCACATCCTAAAACCGACAATTGAAAGTGATCAGCAACACTTTAGGGTATCGCGTG 360

                      Predicted start of thrC translation

                                                    ValGlyGluTyrCysValThrProT
 361  GGCGAAGTCACCTTTTTCAACATATTTGAGACGGTGTGGGGGAGTATTGTGTCACCCCTT 420
                  ribosome binding sequence

      rpIleGlyLeuTyrProTrpThrThrPheArgProArgAspAlaSerArgThrProAlaA
 421  GGATAGGGTTATATCCGTGGACTACATTTCGACCGCGTGATGCCAGCCGTACCCCTGCCC 480

      rgPheSerAspIleLeuLeuGlyGlyLeuAlaProAspGlyGlyLeuTyrLeuProAlaT
 481  GCTTCAGTGATATTTTGCTGGGCGGTCTAGCACCAGACGGCGGCCTGTACCTGCCTGCAA 540
```

FIGURE 4 (CONT'D)

```
     hrTyrProGlnLeuAspAspAlaGlnLeuSerLysTrpArgGluValLeuAlaAsnGluG
541  CCTACCCTCAACTAGATGATGCCCAGCTGAGTAAATGGCGTGAGGTATTAGCCAACGAAG  600

     lyTyrAlaAlaLeuAlaArgGluValIleSerLeuPheValAspAspIleProValGluA
601  GATACGCAGCTTTGGCTCGTGAAGTTATCTCCCTGTTTGTTGATGACATCCCAGTAGAAG  660

     spIleLysAlaIleThrAlaArgAlaTyrThrTyrProLysPheAsnSerGluAspIleV
661  ACATCAAGGCGATCACCGCACGCGCCTACACCTACCCGAAGTTCAACAGCGAAGACATCG  720

     alProValThrGluLeuGluAspAsnIleTyrLeuGlyHisLeuSerGluGlyProThrA
721  TTCCTGTCACCGAACTCGAGGACAACATTTACCTGGGCCACCTTTCCGAACCCGCAACCG  780

     laAlaPheLysAspMetAlaMetGlnLeuLeuGlyGluLeuPheGluTyrGluLeuArgA
781  CTGCATTCAAAGACATGGCCATGCAGCTGCTCGGCGAACTTTTCGAATACGAGCTTCGCC  840

     rgArgAsnGluThrIleAsnIleLeuGlyAlaThrSerGlyAspThrGlySerSerAlaG
841  GCCGCAACGAAACCATCAACATCCTGGGCGCTACCTCTGGCGATACCGGCTCCTCTGCGG  900

     luTyrAlaMetArgGlyArgGluGlyIleArgValPheMetLeuThrProAlaGlyArgM
901  AATACGCCATGCGCGGCCGCGAGGGAATCCGCGTATTCATGCTGACCCCAGCTGGCCGCA  960

     etThrProPheGlnGlnAlaGlnMetPheGlyLeuAspAspProAsnIlePheAsnIleA
961  TGACCCCATTCCAGCAAGCACAGATGTTTGGCCTTGACGATCCAAACATCTTCAACATCG 1020

     laLeuAspGlyValPheAspAspCysGlnAspValValLysAlaValSerAlaAspAlaG
1021 CCCTCGACGGCGTTTTCGACGATTGCCAAGACGTAGTCAAGGCTGTCTCCGCCGACGCAG 1080

     luPheLysLysAspAsnArgIleGlyAlaValAsnSerIleAsnTrpAlaArgLeuMetA
1081 AATTCAAAAAAGACAACCGCATCGGTGCCGTGAACTCCATCAACTGGGCACGCCTTATGG 1140
```

FIGURE 4 (CONT'D)

```
     laGlnValValTyrTyrValSerSerTrpIleArgThrThrThrSerAsnAspGlnLysV
1141 CACAGGTTGTGTACTACGTTTCCTCATGGATCCGCACCACAACCAGCAATGACCAAAAGG  1230

     alSerPheSerValProThrGlyAsnPheGlyAspIleCysAlaGlyHisIleAlaArgG
1201 TCAGCTTCTCCGTACCAACCGGCAACTTCGGTGACATTTGCGCAGGCCACATCGCCCGCC  1260

     lnMetGlyLeuProIleAspArgLeuIleValAlaThrAsnGluAsnAspValLeuAspG
1261 AAATGGGACTTCCCATCGATCGCCTCATCGTGGCCACCAACGAAAACGATGTGCTCGACG  1320

     luPhePheArgThrGlyAspTyrArgValArgSerSerAlaAspThrHisGluThrSerS
1321 AGTTCTTCCGTACCGGCGACTACCGAGTCCGCAGCTCCGCAGACACCCACGAGACCTCCT  1380

     erProSerMetAspIleSerArgAlaSerAsnPheGluArgPheIlePheAspLeuLeuG
1381 CACCTTCGATGGATATCTCCCGCGCCTCCAACTTCGAGCGTTTCATCTTCGACCTGCTCG  1440

     lyArgAspAlaThrArgValAsnAspLeuPheGlyThrGlnValArgGlnGlyGlyPheS
1441 GCCGCGACGCCACCCGCGTCAACGATCTATTTGGTACCCAGGTTCGCCAAGGCGGATTCT  1500

     erLeuAlaAspAspAlaAsnPheGluLysAlaAlaAlaGluTyrGlyPheAlaSerGlyA
1501 CACTGGCTGATGACGCCAACTTTGAGAAGGCTGCAGCAGAATACGGTTTCGCCTCCGGAC  1560

     rgSerThrHisAlaAspArgValAlaThrIleAlaAspValHisSerArgLeuAspValL
1561 GATCCACCCATGCTGACCGTGTGGCAACCATCGCTGACGTGCATTCCCGCCTCGACGTAC  1620

     euIleAspProHisThrAlaAspGlyValHisValAlaArgGlnTrpArgAspGluValA
1621 TAATCGATCCCCACACCGCCGACGGCGTTCACGTGGCACGCCAGTGGAGGGACGAGGTCA  1680

     snThrProIleIleValLeuGluThrAlaLeuProValLysPheAlaAspThrIleValG
1681 ACACCCCAATCATCGTCCTAGAAACTGCACTCCCAGTGAAATTTGCCGACACCATCGTCG  1740
```

FIGURE 4 (CONT'D)

luAlaIleGlyGluAlaProGlnThrProGluArgPheAlaAlaIleMetAspAlaProP
1741 AAGCAATTGGTGAAGCACCTCAAACTCCAGAGCGTTTCGCCGCGATCATGGATGCTCCAT 1800

heLysValSerAspLeuProAsnAspThrAspAlaValLysGlnTyrIleValAspAlaI
1801 TCAAGGTTTCCGACCTACCAAACGACACCGATGCAGTTAAGCAGTACATAGTCGATGCGA 1860

leAlaAsnThrSerValLys
1861 TTGCAAACACTTCCGTGAAGTAACTTGCTTTACGCCAAGGCCTGATTCCTCTCTTTATGG 1920

1921 GATGGAACCAGGCCTTTCGCATTGAGTGGCGTTTTAAGGCCTCCAATTCTTAGAACGGGT 1980
- -
Computer predicted terminator structure, *+termination point

1981 GTTTGACATGGAGGGGTCACAGTCAAGCCGTTAGAAGCGATTCTGGGAGGGCAAGTTTTT 2040

2041 CGGAGTTGGAGGTCGAATTTCCGCTGAACTGATGGGAACCAGACAGGCGTGACAAGATTG 2100

2101 GCTAAAAACCTGAAGTTTTGTCACGCCTGTCTGGTTTCCCTCTTGTCGGTGCGAGCGAGT 2160

2161 CCCTTGAACGACACAGATCGCGCCAAATGGAAGTGTCTGCGACCCCAGAATATTTGATTC 2220

2221 CCCGGTCCGAGTCGTGCGAAAAATGCTCTGGTTAGTCCTCGATCATCGCAATCGCATCAA 2280

2281 TTTCCACAGTTGCACCATAAGGAAGCGATGATGCACCCACGAAAGAGCGTGCCGGGCGGC 2340

2341 CTTCGAGGAAATGCTCTCGGAATTGCTCGTTGCATTCTTCGCGCAGGCTGATGTCGGTGA 2400

2401 CAAAGTAAGTGAGTTTCACAACGTCTTTGAGTTCACCACCAGCGGTCGGAGGCGTTCACG 2460

2461 CATGCGTTCAAGTGCTGCATCAACTGCTTCTTTACGACCGACGACTGGTTGGTAGTCCTT 2520

FIGURE 4 (CONT'D)

2521 GTCTACTGAAAGAGCGCCGGAGACGAAGATGAAATTTCCGACGCGTTTTGCGGGGGACTT 2580

2581 ATGGGTGATCATTCGACATGTGGCCAACCATAGCTGTTTCCCCGAAGAGAGTGCCGGAAC 2640

2641 AGGCATTTTAGAGGTGGGGGAGCACTTCTTCGTAAATCTGGGTCAGTACTTCGCTTGCTG 2700

2701 GTCGGCGCTGGATGTTGAAGATGACGTGGTCGATGCCAAGTTCCGAAAGCGGTGGAGGTC 2760

2761 TTCGATGAGTTCCTGGCTGCCTACCTCTACGCCAGAGTGATTTCTTTGTGGGTGTTTCCT 2820

2821 TCGGTGAGGTTGAGCCCCATGGAGGAAATCAACAAGGGGCGGGTGCCACCACGGGCTTTG 2880

2881 TCCCAGAGATCGAGGCGTCCGACTTGAGCTTCAGCGGGGCGGTAGTAGGTTGCCCATCCG 2940

2941 TCGGCGTTTCGGGCGATCCATTGCACTGTTTGTCGGGCAGAACCTACAGCGATCATGGGG 3000

3001 ATCTGAGCTTCAGGTGGCGTGGTTGGCGCAAATTCAAGGTCGGCCCGCATCGCAGGATCC 3060

3061 TTCGACAAAGCTGCACGCAAAATTGCCCACCCAGACTGAATATCAGCGCGTCGATTGTCT 3120

3121 AAGCTTTTCGGAAAAATCTCGAATTC 3146

FIGURE 5A

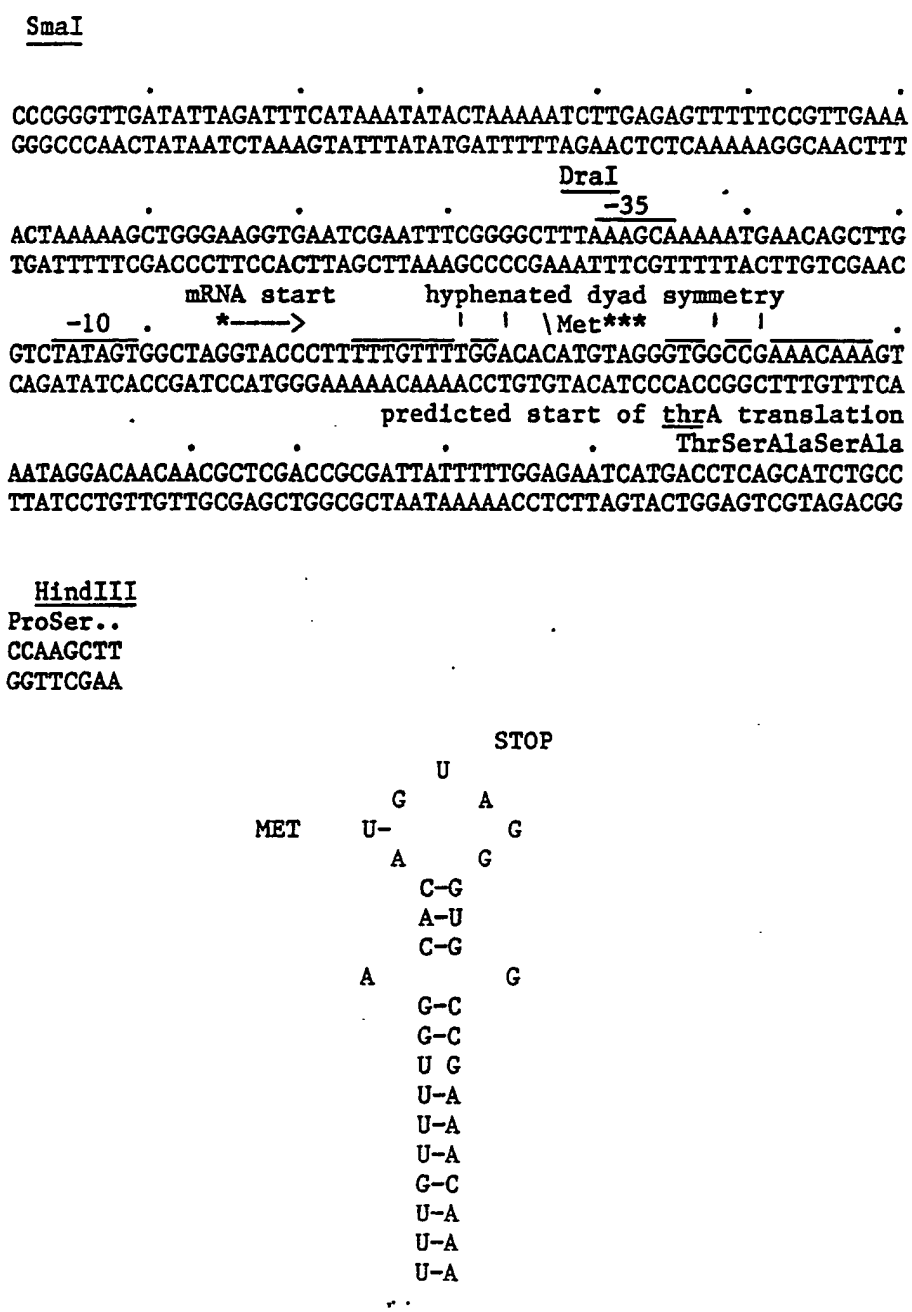
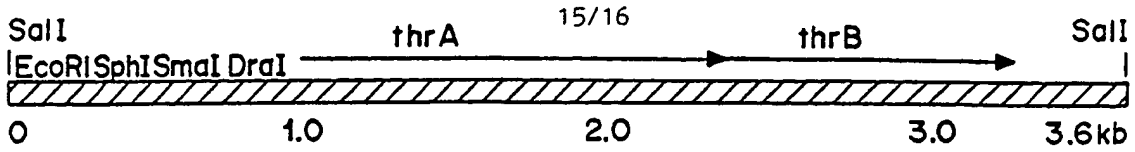


FIGURE 5B

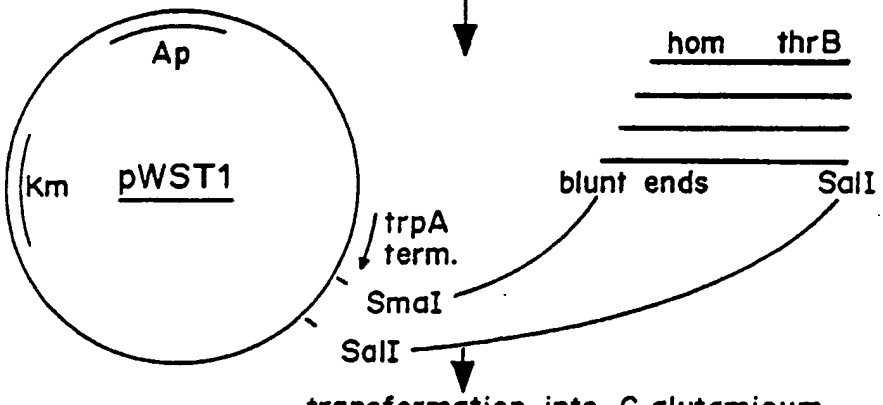


1.) determination of the approximate location of the thrAB promoter

- EcoRI, SphI, SmaI, or DraI restriction
- make blunt ends if necessary
- SalI restriction
- Isolate hom-thrB containing fragment
- ligate into pWST1 cut with SmaI and SalI

2.) fine deletion mapping of the thrAB promoter with Bal31

- SmaI restriction
- digest with Bal31
- SalI restriction
- Isolate hom-thrB containing fragment
- ligate into pWST1 cut with SmaI and SalI



transformation into C.glutamicum thr/met auxotroph AS253

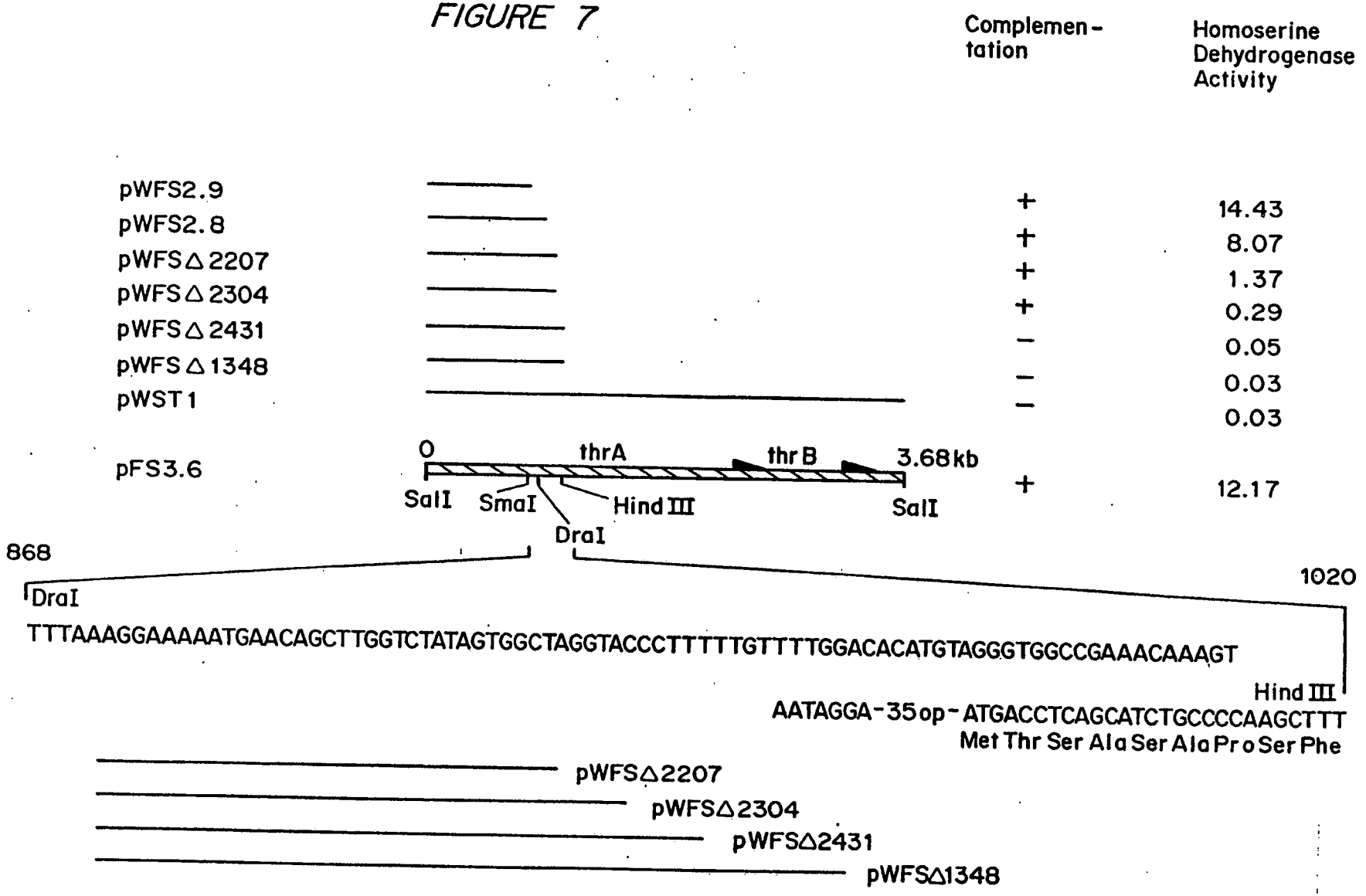
complementation analysis

homoserine dehydrogenase enzyme assays

determination of nucleotide sequences

FIGURE 6

FIGURE 7



PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 :
C12N 15/00, C12P 13/08
C12N 9/04, 9/12, 9/88

A3

(11) International Publicati n Number: WO 88/ 09819

(43) International Publicati n Date:
15 December 1988 (15.12.88)

(21) International Application Number: PCT/US88/02029

(22) International Filing Date: 10 June 1988 (10.06.88)

(31) Priority Application Number: 062,552

(32) Priority Date: 12 June 1987 (12.06.87)

(33) Priority Country: US

(71) Applicant: MASSACHUSETTS INSTITUTE OF TECHNOLOGY [US/US]; 77 Massachusetts Avenue, Cambridge, MA 02139 (US).

(72) Inventors: SINSKEY, Anthony, J. ; 285 Commonwealth Avenue, Boston, MA 02115 (US). FOLLETTIE, Maximillian, T. ; 474 Memorial Drive, Cambridge, MA 02139 (US). LIEBL, Wolfgang ; Hart Str. 17, D-8057 Eching (DE). PEOPLES, Oliver, P. ; 104 Fairmont Street, Arlington, MA 02174 (US).

(74) Agent: PABST, Patrea, L.; Kilpatrick & Cody, 100 Peachtree Street, Suite 3100, Atlanta, GA 30043 (US).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

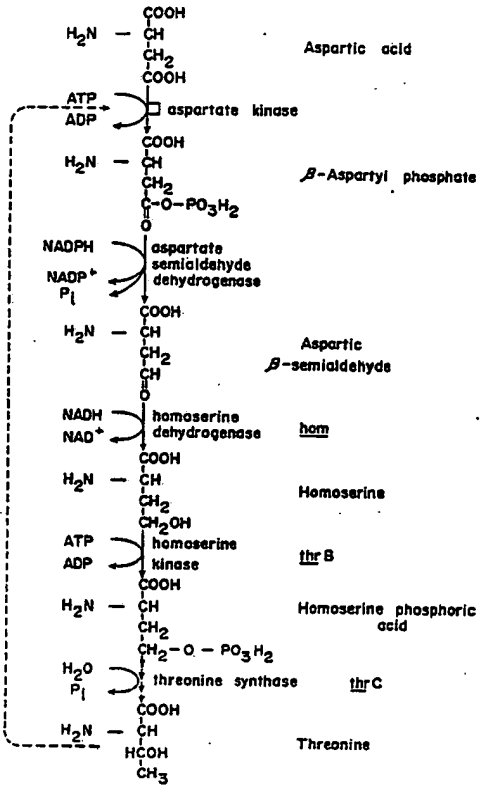
Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
9 March 1989 (09.03.89)

(54) Title: C. GLUTAMICUM THREONINE BIOSYNTHETIC PATHWAY

(57) Abstract

The present invention is a method for the isolation and characterization of *C. glutamicum* genes involved in amino acid biosynthesis, specifically, encoding *hom*, *thr*B, and *thr*C, and sequences regulating their expression. Techniques for modifying or replacing these sequences and means for facilitating further isolations and characterizations, including promotor probe vectors which are useful in screening for high efficiency and regulatable promoters and repressors, are also disclosed. A *C. glutamicum* genomic library was constructed by cleaving chromosomal DNA with restriction enzymes, inserting the DNA fragments into an appropriate vector, and transforming the resulting recombinant molecules (rDNA) into *C. glutamicum*. Amino acid biosynthetic genes *hom*, *thr*B, and *thr*C, encoding homoserine dehydrogenase, homoserine kinase, and threonine synthetase, respectively, were isolated by complementation of *C. glutamicum* auxotrophs. The *hom-thr*B genes were subcloned on a 3.6 kb *Sal*I generated chromosomal fragment while *thr*C activity was isolated from a second recombinant plasmid within the genomic library and subcloned on a 2.7 kb *Sph*I generated fragment. The *hom-thr*B and *thr*C loci, and regulatory sequences, were identified by enzyme assays, complementation of defined *E. coli* auxotrophs, S1 nuclease and deletion mapping.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	ML	Mali
AU	Australia	GA	Gabon	MR	Mauritania
BB	Barbados	GB	United Kingdom	MW	Malawi
BE	Belgium	HU	Hungary	NL	Netherlands
BG	Bulgaria	IT	Italy	NO	Norway
BJ	Benin	JP	Japan	RO	Romania
BR	Brazil	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SN	Senegal
CH	Switzerland	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
DE	Germany, Federal Republic of	LU	Luxembourg	TG	Togo
DK	Denmark	MC	Monaco	US	United States of America
FI	Finland	MG	Madagascar		

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 88/02029

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[4]: C 12 N 15/00; C 12 P 13/08; C 12 N 9/04; C 12 N 9/12; C 12 N 9/88

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[4]	C 12 N; C 12 P

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [10]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
P,X	EP, A, 0233581 (KYOWA HAKKO KOGYO CO. LTD) 26 August 1987 see the whole document	1-3,7,9, 11,18-20, 22-27
X	EP, A, 0204326 (KYOWA HAKKO KOGYO CO.) 10 December 1986 see the whole document	1-3,7,11, 22-24,26
X	EP, A, 0179338 (AJINOMOTO CO.) 30 April 1986 see page 21, line 19 - page 26; page 40, line 1 - page 47; claims	1-3,11,23, 24
X	WO, A, 87/02984 (SIMPSON) 21 May 1987 see example 10; claim 1	1,2,4-7, 9,12-15, 18-21,23-25
X	EP, A, 0088166 (KYOWA HAKKO CO.) 14 September 1983 see example 2	1,2,7,9,11
A	US, A, 4278765 (V.G. DEBABOV) ./.	

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
19th January 1989	07. 02. 89
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	M. VAN MOL

Form PCT/ISA/210 (second sheet) (January 1985)

International Application No. PCT/US 88/02029

III. DOCUMENTS CONSIDERED T BE RELEVANT (C NTINUED FR M THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
	14 July 1981	
A	EP, A, 0066129 (AJINOMOTO CO.) 8 December 1982	
A	EP, A, 0190921 (ENGENICS) 13 August 1986	
Y	EP, A, 0215388 (AJINOMOTO CO.) 25 March 1987 see examples 1,2	16-17
Y	Journal of Biotechnology, volume 5, no. 4, May 1987, Elsevier Science Publishers B.V. (Biomedical Division), (Amsterdam, NL), Y. Morinaga et al.: "Expression of Escherichia coli promoters in Brevibacterium lactofermentum using the shuttle vector pEB003", pages 305-312	16-17
Y	Food Technology, volume 40, no. 10, October 1986, Institute of Food Technologists, (Chicago, Illinois, US), M.T. Follettie et al.: "Recombinant DNA technology for Corynebacterium glutamicum", pages 88-94 see the whole document	16-17
Y	Chemical Abstracts, volume 101, no. 21, 19 November 1984, (Columbus, Ohio, US), K. Yoshimura et al.: "Construction of a promoter-probe vector for a Bacillus subtilis host by using the trpD$^+$ gene of Bacillus amyloliquefaciens", see page 176, abstract 185045p, & J. Bacteriol. 1984, 159(3), 905-12	16-17

FURTHER INF RMATION CONTINUED FR M THE SECOND SHEET

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers....., because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim numbers....., because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim numbers....., because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☐ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

Claims 1-15,18-28
Claims 16,17
For further information please see Form PCT/ISA/206
dated 18th October 1988.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest.

☒ No protest accompanied the payment of additional search fees.

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. US 8802029 SA 23095

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 24/01/89
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A- 0233581	26-08-87	JP-A- 62186795	15-08-87
EP-A- 0204326	10-12-86	JP-A- 62091193	25-04-87
		JP-A- 62232392	12-10-87
EP-A- 0179338	30-04-86	JP-A- 61088889	07-05-86
WO-A- 8702984	21-05-87	AU-A- 6737287	02-06-87
		EP-A- 0245497	19-11-87
		JP-T- 63501687	14-07-88
EP-A- 0088166	14-09-83	JP-A- 58126789	28-07-83
		AU-A- 9191982	07-07-83
		CA-A- 1202257	25-03-86
		AU-B- 557228	11-12-86
US-A- 4278765	14-07-81	None	
EP-A- 0066129	08-12-82	JP-A- 57186496	16-11-82
		US-A- 4452890	05-06-84
EP-A- 0190921	13-08-86	JP-A- 61247392	04-11-86
EP-A- 0215388	25-03-87	JP-A- 62151184	06-07-87